

GenCore version 5.1.6

Run on: October 27, 2003, 14:53:57 ; Search time 84 Seconds
(without alignments)
1105.417 Million cells up

perfect score: 3103

Scoring table: BLOSUM62

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 19Jun03:*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %			DB	ID	Description
		Match	Length	DB			
1	3103	100.0	585	10	AAP90388	Mature human serum	
2	3103	100.0	595	11	AAR05318	Human serum albumi	
3	3103	100.0	585	11	AAP08457	Human serum albumi	
4	3103	100.0	585	16	AAR80301	Human serum albumi	
5	3103	100.0	585	18	AAO20111	PSA protein; sequen	
6	3103	100.0	585	21	AAV94873	Amnic acid sequenc	
7	3103	100.0	595	21	AAV83946	Yeast codon-biased	
8	3103	100.0	585	22	ABB79096	Human mature album	
9	3103	100.0	585	22	AAE13399	Human albumin (HA)	

10	3103	100.0	585	22	AAW52567	Mature human serum
11	3103	100.0	585	22	AAE13129	Human albumin (HA)
12	3103	100.0	585	22	AAE13403	Human albumin (HA)
13	3103	100.0	585	22	AAE09578	Human serum albumi
14	3103	100.0	585	23	ABG71291	Glycosylated protei
15	3103	100.0	585	23	ABG63321	Human serum albumi
16	3103	100.0	585	23	ABJ00986	B lymphocyte stimu
17	3103	100.0	585	23	ABG33847	Human B lymphocyte
18	3103	100.0	585	23	AAU75220	Mature form of hum
19	3103	100.0	609	21	AAAB36542	Recombinant human
20	3103	100.0	609	21	AAAB36549	Recombinant humar
21	3103	100.0	609	21	AAV78147	Pre human serum al
22	3103	100.0	609	24	ABU57252	Human serum albumi
23	3103	100.0	609	24	ABU57253	Human serum albumi
24	3103	100.0	610	14	AAAR39510	Chimeric human ser
25	3103	100.0	616	24	AAE30916	Val18-GLP-1-human s
26	3103	100.0	624	24	AAE30919	Human serum albumi
27	3103	100.0	631	24	AAE30917	Val18-GLP-1-linker
28	3103	100.0	640	24	AAE30918	Gly8-GLu22-GLP-1-C
29	3103	100.0	640	24	AAE30920	Extendin-4-linker-b
30	3103	100.0	670	21	AAAB36543	Recombinant human
31	3103	100.0	670	21	AAAB36550	Recombinant human
32	3103	100.0	783	14	AAAR39473	Prepro-HSA-G-CSP C
33	3103	100.0	787	14	AAAR39477	G-CSP- (Gly)4-HSA C
34	3103	100.0	853	14	AAAR39472	HSA-VWF (470-713) f
35	3099	99.9	585	10	AAAP93344	Sequence of mature
36	3099	99.9	585	19	AAW59841	Mature protein of
37	3099	99.9	608	17	AAAP96229	Human serum albumi
38	3099	99.9	609	4	AAAP96229	Sequence of human
39	3099	99.9	609	17	AAAP96232	Human serum albumi
40	3099	99.9	609	17	AAAR94572	Cancer metastasis
41	3099	99.9	609	17	AAAB88913	Human serum albumi
42	3099	99.9	609	19	AAAB48095	Human serum albumi
43	3099	99.9	609	20	AAAY06994	Human albumin. Ho
44	3099	99.9	609	22	AAAB04148	Myosin light chain
45	3099	99.9	609	23	ABG32802	Human serum albumi

ALIGNMENTS

RESULT 1	
AAP90388	AAP90388 standard; protein; 585 AA.
XX	
XX	
XX	AAP90388;
XX	
XX	25-MAR-2003 (updated)
XX	01-NOV-1989 (first entry)
XX	
XX	Mature human serum albumin polypeptide.
XX	
XX	Human serum albumin; mature protein; new polypeptides;
XX	plasma expanders.
XX	
XX	Homo sapiens (Human).
XX	
XX	EP322094-A.
XX	
XX	28-JUN-1989.
XX	
XX	25-OCT-1998; 88EP-0310000.
XX	
XX	30-OCT-1987; 87GB-0025529.
XX	
XX	{DELZ } DELTA BIOTECHNOLOGY LTD.
XX	
XX	Bailance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX	
XX	WPI; 1989-186464/26.
XX	N-PSDB; AAN90128.
XX	
XX	New N-terminal fragments of human serum albumin
XX	

P - esp. useful as blood plasma expanders.

XX
XX
PS Disclosure; fig 2; 20pp; English.
XX
XX Mature protein of human serum albumin (see corresp. AAN90128).
CC Used to make new N-terminal fragments which are used as plasma
CC expanders, or as substitutes for HSA or BSA, in tissue culture
CC media.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120

QY 121 DVMCTAFHDNEETFKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180
DB 121 DVMCTAFHDNEETFKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180

QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFVSKLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFVSKLVTDLT 240

QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEYPA 300
DB 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEYPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360

QY 361 CAADAPHECYAKVDFEKPPLVEEPQNLIKNCSELFQOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADAPHECYAKVDFEKPPLVEEPQNLIKNCSELFQOLGEYKFNALLVRYTKKVPQVST 420

QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEEDYLSVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEEDYLSVLNQLCVLHEKTPVSDRVTCKCTES 480

QY 481 LVNRRPCFSALVEDETYPVPKEFNAETTFTHADICTLSEKERQIKKOTALVELVKKPKAT 540
DB 481 LVNRRPCFSALVEDETYPVPKEFNAETTFTHADICTLSEKERQIKKOTALVELVKKPKAT 540

QY 541 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 2

AAR05318

ID AAR05318 standard; protein; 585 AA.

XX

AC AAR05318;

XX

DT 08-OCT-1990 (first entry)

XX Human serum albumin gene product.

XX

XX Human serum albumin; HSA-A; yeast; ds.

XX

OS Homo sapiens.

XX

PN JP02117384-A.

XX

PD 01-MAY-1990.

XX

PF 26-OCT-1988; 88JP-0268302.

XX

PR 26-OCT-1988; 88JP-0268302.

XX

PA (TOFU) TOA NENRYO KOGYO KK.

XX

KPI; 1990-176228/23.

DR N-FSDB; AAO04719.

XX

Human serum albumin prepn. by yeast host -

PT by culturing transformed plasmid yeast to produce serum, and

PT removing it.

XX

PS Disclosure; Page ?; ?pp; Japanese.

XX

Mature HSA-A may be produced using the sequence incorporated into a

CC plasmid vector with suitable controllers, and transferred to a yeast

CC expression system.

XX

SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAE 60

DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120

QY 121 DVMCTAFHDNEETFKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180

DB 121 DVMCTAFHDNEETFKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180

QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFVSKLVTDLT 240

DB 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFVSKLVTDLT 240

QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEYPA 300

DB 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEYPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360

DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360

QY 361 CAADAPHECYAKVDFEKPPLVEEPQNLIKNCSELFQOLGEYKFNALLVRYTKKVPQVST 420

DB 361 CAADAPHECYAKVDFEKPPLVEEPQNLIKNCSELFQOLGEYKFNALLVRYTKKVPQVST 420

QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEEDYLSVLNQLCVLHEKTPVSDRVTCKCTES 480

DB 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEEDYLSVLNQLCVLHEKTPVSDRVTCKCTES 480

QY 481 LVNRRPCFSALVEDETYPVPKEFNAETTFTHADICTLSEKERQIKKOTALVELVKKPKAT 540

DB 481 LVNRRPCFSALVEDETYPVPKEFNAETTFTHADICTLSEKERQIKKOTALVELVKKPKAT 540

QY 541 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKKLVAASQAALGL 585

DB 541 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 3

AAR08457

ID AAR08457 standard; protein; 585 AA.

XX

AC AAR08457;

XX 25-MAR-2003 (updated)
 DT 16-APR-1991 (first entry)
 XX Human serum albumin.
 DE HSA: folding; ss.
 KW HSA: folding; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Key
 FH Region 123..303
 FT /label= A
 FT Region 1..303
 FT /label= B
 FT Region 123..585
 FT /label= C
 XX JP02227079-A.
 PN 10-SEP-1990.
 PD 25-AUG-1989; 89JP-0217540.
 XX 25-AUG-1989; 89JP-0217540.
 XX (TOFU) TONEN CORP.
 PA WPI; 1990-317325/42.
 DR N-PSDB; AAQ06099.
 XX New human serum albumin fragments - used to bond medicines and for
 PT stable folding of protein(s).
 XX Claim 1; Fig 8; 24pp; Japanese.
 XX Fragments A-C of HSA are expressed as fusion proteins with the
 CC signal peptide of E. coli alkaline phosphatase. The fragments are
 CC selected for their specific properties. The C-terminal truncated
 CC fragment, B, does not bind long-chain fatty acids but does bind to
 CC various medicines at the central region. The N-terminal truncated
 CC fragment, C, has good stability in protein folding. The central
 CC segment, A, has characteristics of both B and C.
 CC See also AAQ06096-Q06098.
 CC (Updated on 25-MAR-2003 to correct PD field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX Sequence 585 AA;
 SQ
 Query Match 100.0%; Score 3103; DB 11; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKSEVAHRPKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSAE 60
 DB 1 DAHKSEVAHRPKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSAE 60
 QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
 DB 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
 QY 121 DVMCTAFHDNETFLKKYLYEYARSHPVFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
 DB 121 DVMCTAFHDNETFLKKYLYEYARSHPVFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
 QY 181 KLDELDEGKASSAKORUKCASLQKFGERAFKAWAVARLSORFPKAEFAEYSKLVTLTK 240
 DB 181 KLDELDEGKASSAKORUKCASLQKFGERAFKAWAVARLSORFPKAEFAEYSKLVTLTK 240
 QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLUKCECEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLUKCECEKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVWILLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVWILLRLAKTYETTLK 360
 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
 DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
 QY 421 PTLVEVSRNLGVGSKCCCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGVGSKCCCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
 QY 541 KEOLKAVMDDFAAVFEKCKCKADDKETCFAEKGKLVAAASQAALGL 585
 DB 541 KEOLKAVMDDFAAVFEKCKCKADDKETCFAEKGKLVAAASQAALGL 585
 RESULT 4
 AAR8C301
 ID AAR80301 standard; Protein; 585 AA.
 XX AAR80301;
 AC XX
 XX 25-MAR-2003 (updated)
 DT 17-JAN-1996 (first entry)
 XX Human serum albumin.
 DE Serum albumin; HSA; aspartyl protease-3; Yap3p;
 XX Saccharomyces cerevisiae.
 KW Homo sapiens.
 OS WO9523857-A1.
 XX 08-SEP-1995.
 XX 01-MAR-1995; 95WO-GB00434.
 XX 03-MAR-1994; 94GB-0004270.
 XX (DELZ) DELTA BIOTECHNOLOGY LTD.
 XX KerryWilliams SM, Gilbert SC;
 XX WPI; 1995-320572/41.
 DR N-PSDB; AAQ98695.
 XX Yeast with reduced levels of aspartyl protease 3 proteolytic
 PT activity - used to secrete human albumin without prodrn. of the 45
 PT kD fragment
 XX Example 1; Page 26-28; 50pp; English.
 XX The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected
 CC to site-directed mutagenesis to investigate the role of
 CC endoproteases in the generation of a 45 kDa albumin fragment: obtd.
 CC when the cDNA is expressed in S. cerevisiae. Mutations were: R410A;
 CC L407A, L408V, V409A; and R410A, K413Q, K414Q. The latter set of
 CC mutations, especially, improved stability of HSA to yeast Yap3p
 CC proteolytic cleavage, allowing increased prodrn. of recombinant HSA.
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX Sequence 585 AA;
 SQ
 Query Match 100.0%; Score 3103; DB 16; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 DAHSEVAHRFKDGBENFKALVLIATAFYLOQCPECHVKLVNEVTEFAKTCVADESAAE 60
Db      1 DAHSEVAHRFKDGBENFKALVLIATAFYLOQCPECHVKLVNEVTEFAKTCVADESAAE 60
QY      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
Db      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
QY      121 DMCTAFHNDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db      121 DMCTAFHNDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
QY      181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORPFAEVSJKVTDLTJK 240
Db      181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORPFAEVSJKVTDLTJK 240
QY      241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPELLEKSHHCIAEVENDEMPA 300
Db      241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPELLEKSHHCIAEVENDEMPA 300
QY      301 DLPSLAADRVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db      301 DLPSLAADRVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360
QY      361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQGEYKFNQALLVRYTKKVPQVST 420
Db      361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQGEYKFNQALLVRYTKKVPQVST 420
QY      421 PTLVEVSRLNGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db      421 PTLVEVSRLNGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
QY      481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
Db      481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
QY      541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEKGKLVAAASQAALGL 585
Db      541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEKGKLVAAASQAALGL 585

```

RESULT 5

AA020111
ID AAC20111 standard; Protein; 585 AA.

XX AAC20111;

XX AAC20111;

XX 06-AUG-2002 (first entry)

XX HSA protein sequence related to the growth hormone protein.

XX Serum albumin-growth hormone fusion protein; growth hormone;

XX Down's syndrome.

XX Unidentified.

XX KR99076789-A.

XX 15-OCT-1999.

XX 25-JUN-1998; 98KR-0704914.

XX 30-DEC-1995; 95GB-0026733.

XX 19-DEC-1996; 96WO-GB03164.

XX (DELZ) DELTA BIOTECHNOLOGY LTD.

XX WPI; 1997-363680/55.

XX N-PSDB; AAK99568.

XX Serum albumin-growth hormone fusion protein - useful to treat growth

XX hormone related diseases, e.g. Down's syndrome

XX Disclosure; Fig 6; 21pp; Korean.
XX The invention relates to a serum albumin-growth hormone fusion protein -
CC useful to treat growth hormone related diseases such as Down's syndrome.
CC This sequence represents a HSA protein related to the serum albumin-
CC growth hormone protein of the invention.
XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 18; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 DAHSEVAHRFKDGBENFKALVLIATAFYLOQCPECHVKLVNEVTEFAKTCVADESAAE 60
Db      1 DAHSEVAHRFKDGBENFKALVLIATAFYLOQCPECHVKLVNEVTEFAKTCVADESAAE 60
QY      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
Db      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
QY      121 DMCTAFHNDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db      121 DMCTAFHNDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
QY      181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORPFAEVSJKVTDLTJK 240
Db      181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORPFAEVSJKVTDLTJK 240
QY      241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPELLEKSHHCIAEVENDEMPA 300
Db      241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPELLEKSHHCIAEVENDEMPA 300
QY      301 DLPSLAADRVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db      301 DLPSLAADRVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360
QY      361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQGEYKFNQALLVRYTKKVPQVST 420
Db      361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQGEYKFNQALLVRYTKKVPQVST 420
QY      421 PTLVEVSRLNGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db      421 PTLVEVSRLNGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
QY      481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
Db      481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
QY      541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEKGKLVAAASQAALGL 585
Db      541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEKGKLVAAASQAALGL 585

```

RESULT 6

AA020111
ID AAC20111 standard; Protein; 585 AA.

XX AAC20111;

XX AAC20111;

XX 08-AUG-2000 (first entry)

XX Amino acid sequence of a human albumin protein.

XX Human; albumin; ischemic state; serum protein; metal ion salt;

XX peroperative ischemia; ischemia; myocardial infarction;

XX progressive coronary artery disease.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 1

/note= "optionally acetylated, and claimed under
claim 56"

PT WO200020840-A1.
XX 13-APR-2000.
XX 01-OCT-1999; 99WO-US22905.
XX 02-OCT-1998; 98US-0102738.
PR 02-OCT-1998; 98US-016581.
PR 02-OCT-1998; 98US-0165926.
PR 11-JAN-1999; 99US-0115392.
XX (ISCH-) ISCHEMIA TECHNOLOGIES INC.
PA Bar-Or D, Lau E, Winkler JV;
XX WPI; 2000-303843/26.
XX New method for the continuous detection of ischemic states comprises
PT detecting and quantifying the existence of an alteration of the serum
PT protein albumin.
XX Disclosure; Page 97-100; 105pp; English.
XX The present sequence represents human albumin protein. The specification
CC describes a method for the continuous detection of ischemic states. The
CC method comprises detecting and quantifying the existence of an alteration
CC of the serum protein albumin. The method comprises contacting a
CC biological sample containing albumin from the patient with an excess
CC quantity of a metal ion salt, where the metal ion binds to the N-terminus
CC of naturally occurring human albumin, to form a mixture containing bound
CC metal ions and unbound metal ions, and then determining the amount of
CC metal ions bound to the albumin N-terminus. The amount of bound metal
CC ions is correlated to a known value to determine the occurrence or
CC non-occurrence of an ischemic event. The methods are useful for detection
CC of ischemic states. The methods are also useful for distinguishing
CC peroperative ischemia from ischemia caused by, amongst other things,
CC myocardial infarctions and progressive coronary artery disease.
XX
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 21; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDGBENFKALVLIAPFAQYLOQCPEFHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDGBENFKALVLIAPFAQYLOQCPEFHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLQHKDQNPILPRIVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLQHKDQNPILPRIVRPEV 120
QY 121 DVNCTAFHNDNEETFLKYLIEARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
DB 121 DVNCTAFHNDNEETFLKYLIEARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
QY 181 KLDELRLDEGKASAKORLJKASLQKQGERAFKAWAVARLSQRFPAEVAESKLVTDITK 240
DB 181 KLDELRLDEGKASAKORLJKASLQKQGERAFKAWAVARLSQRFPAEVAESKLVTDITK 240
QY 241 VHTTECHGDLLECAADRADLAKYICENQDISISKLEKCECEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECAADRADLAKYICENQDISISKLEKCECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYVSWLLRLAKTYETITLKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYVSWLLRLAKTYETITLKC 360
QY 361 CAAADPHECYAKVDFEFPKLVPEEPQNLKONCELFEQLGKEYKFNALLVRYTKKVPQVST 420
DB 361 CAAADPHECYAKVDFEFPKLVPEEPQNLKONCELFEQLGKEYKFNALLVRYTKKVPQVST 420

DB 361 CAAADPHECYAKVDFEFPKLVPEEPQNLKONCELFEQLGKEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHAD:CTLSEKERQIKKOTALVELVKKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHAD:CTLSEKERQIKKOTALVELVKKHKPKAT 540
QY 541 KEOLKAVMDDFAAAFVEKCKCKADDDKTCFAEGEGKKLVAASQAALGL 585
DB 541 KEOLKAVMDDFAAAFVEKCKCKADDDKTCFAEGEGKKLVAASQAALGL 585
RESULT 7
AAAY83946
ID AAAY83946 standard; Protein; 585 AA.
XX AC AAAY83946;
XX DT 28-JUL-2000 (first entry);
XX DE Yeast codon-biased recombinant human serum albumin protein.
XX KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
XX over-lapping oligonucleotide; expression vector.
XX OS Homo sapiens.
XX PN CNI2339103-A.
XX PD 22-DEC-1999.
XX PF 17-JUN-1998; 98CN-0102506.
XX PR 17-JUN-1998; 98CN-0102506.
XX PA (HAIJ-) HAIJI BIOENGINEERING CO LTD.
XX PI Li S, Lu D;
XX WPI; 2000-351198/31.
XX N-PSDB; AAA10091.
XX Process for preparing recombinant human serum albumin - which comprises
PT yeast biased sex codons
XX PS Disclosure; Fig 1; 44pp; Chinese.
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to
CC comprise a yeast codon bias. The complete HSA gene (AAA10091) was
CC generated as three synthetic fragments (AAA10092-A10094) joined by
CC recombinant DNA technology. Each HSA fragment was synthesised from
CC overlapping oligonucleotide fragments that were extended. This sequence
CC represents the complete sequence of the HSA encoded by the human gene
CC with a yeast codon bias. The invention also covers a recombinant
CC expression vector, yeast host cells carrying the recombinant expression
CC vector and the process for producing human serum albumin in the yeast
CC host cell, especially in secretory mode.
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 21; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDGBENFKALVLIAPFAQYLOQCPEFHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDGBENFKALVLIAPFAQYLOQCPEFHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCXSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDXDDNPNIPLVVRPEV 120
Db 61 NCXSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDXDDNPNIPLVVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKQBLKASLOKQGERAFKAWAVARLSQREPKAEFAEVSXKLVDTLK 240
Db 181 KLDELDEGKASSAKQBLKASLOKQGERAFKAWAVARLSQREPKAEFAEVSXKLVDTLK 240
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKJLKECEKXPJLEKSHCIAEVENDEMPA 300
Db 241 VHTCCHGDLLECADRADLAKYICENQDSISSKJLKECEKXPJLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRJAKTYVETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRJAKTYVETTLK 360
QY 361 CAAADPHECYAKVDFEKLVEEPONLIKONCELFEQLGEYKFONALLVRYTKKYPQVST 420
Db 361 CAAADPHECYAKVDFEKLVEEPONLIKONCELFEQLGEYKFONALLVRYTKKYPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVNLQCLVHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVNLQCLVHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVYKFEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDVETVYKFEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
QY 541 KEQLKAVMDFAAFVEKCKADKQETCFABEGKGLVAASQAALGL 585
Db 541 KEQLKAVMDFAAFVEKCKADKQETCFABEGKGLVAASQAALGL 585

RESULT 8
ABB79006

ID ABB79006 standard; Protein: 585 AA.

AC ABB79006;

XX 01-AUG-2002 (first entry)

XX Human mature albumin protein SEQ ID NO:18.

XX Human; growth hormone; hGH; albumin; human serum albumin; HSA;
KW albumin fusion protein; cytosolic; anorectic; immunosuppressive;
KW antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;
KW non-Hodgkin's lymphoma; obesity; type I diabetes mellitus; rheumatoid arthritis;
KW type I diabetes mellitus; rheumatoid arthritis.

XX Homo sapiens.

PH Key Location/Qualifiers
FT Domain 1..194
FT /label= 1
FT Domain 1..105
FT /label= subdomain
FT Disulfide-bond 53..62
FT Disulfide-bond 75..91
FT Disulfide-bond 90..101
FT Region 106..119
FT /note= "flexible inter-subdomain linker region"
FT Domain 120..194
FT /label= subdomain
FT Disulfide-bond 124..169
FT Disulfide-bond 168..177
FT Domain 195..387
FT /label= 2
FT Domain 195..291
FT /label= subdomain

FT Disulfide-bond 245..253
FT Disulfide-bond 265..279
FT Disulfide-bond 278..289
FT Region 292..315
FT /note= "flexible inter-subdomain linker region"
FT Domain 316..387
FT /label= subdomain
FT Disulfide-bond 316..361
FT Disulfide-bond 360..369
FT Domain 388..585
FT /label= 3
FT Domain 388..491
FT /label= subdomain
FT Disulfide-bond 392..438
FT Disulfide-bond 437..448
FT Disulfide-bond 461..477
FT Disulfide-bond 476..487
FT Region 492..511
FT /note= "flexible inter-subdomain linker region"
FT Domain 512..585
FT /label= subdomain
FT Disulfide-bond 514..559
FT Disulfide-bond 558..567
XX WO200179442-A2.
XX 25-OCT-2001.
XX 12-APR-2001; 2001WO-US11850.
XX 12-APR-2000; 2000US-229358P.
XX 25-APR-2000; 2000US-199384P.
XX 21-DEC-2000; 2000US-256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Haseltine WA;
XX WPI; 2001-611723/70.
XX N-PSDB; ABB79006.
XX New albumin fusion proteins, useful for treating diseases and disorders
XX such as cancer, comprise therapeutic protein fused to albumin -
XX Claim 1; Fig 11; 413pp; English.
XX The present invention describes an albumin fusion protein (I) comprising
XX a therapeutic protein: X and a fragment or variant of albumin
XX comprising a the fully defined sequence in ABB79006 of 585 amino acids,
XX (where the fragment or variant has albumin or therapeutic protein: X
XX activity). (I) can have cytostatic, anorectic, immunosuppressive,
XX antidiabetic, antirheumatic, antiarthritic and psoriatic activities.
XX Albumin fusion proteins are stabilised therapeutic proteins e.g.
XX antibodies to C5, C242 and CD80 useful for treating various diseases
XX and disorders such as non-Hodgkin's lymphoma, cancer, obesity,
XX transplant rejection, type I diabetes mellitus, rheumatoid arthritis
XX and psoriasis. Fusing albumin to therapeutic proteins stabilises the
XX therapeutic protein, extends the shelf life and retains the in vitro or
XX in vivo biological activity. It also reduces the need to formulate
XX protein solutions with large excesses of carrier proteins to prevent
XX loss of therapeutic proteins due to factors such as binding to the
XX container. The fusion proteins are easily dispensed with a simple
XX formulation requiring minimal post storage manipulation. The fusion of
XX therapeutic proteins to albumin confers stability in aqueous or other
XX solution. The present sequence represents the mature human albumin (HA)
XX protein which is used in the exemplification of the present invention.

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 22; Length 585;

Best Local Similarity 100.0%; Pred. No. 9 5e-255;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 DVNCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCASLOKGEGERAFKAWAVARLSORPPKAEFAEVSJKLVTDLT 240
 Db 181 KLDELDEGKASSAKORLKCASLOKGEGERAFKAWAVARLSORPPKAEFAEVSJKLVTDLT 240
 QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPA 300
 Db 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPA 300
 QY 301 DPLSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYISVLLRLRAKTYETTTLEKC 360
 Db 301 DPLSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYISVLLRLRAKTYETTTLEKC 360
 QY 361 CAADDPHECYAKVDFSEFKPLVEEPONLIKONCELFQOLGEYKFORALLVRYTKVPQVST 420
 Db 361 CAADDPHECYAKVDFSEFKPLVEEPONLIKONCELFQOLGEYKFORALLVRYTKVPQVST 420
 QY 421 PTLVEVSRLGKVGSKCKHPKAEKRYPCAEEDYLSVVLNQLCVLHEKTPVSDRVTXCCTES 480
 Db 421 PTLVEVSRLGKVGSKCKHPKAEKRYPCAEEDYLSVVLNQLCVLHEKTPVSDRVTXCCTES 480
 QY 481 LVNRRPCFSALEVDVETVVPKFEFNAETTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540
 Db 481 LVNRRPCFSALEVDVETVVPKFEFNAETTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540
 QY 541 KEOLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
 Db 541 KEOLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585

RESULT 10

AAW52567
 ID AAW52567 standard; Protein; 585 AA.

XX AAW52567;

XX 05-FEB-2002 (first entry)

DE Mature human serum albumin.

XX Human; serum albumin; HA; anti-inflammatory; immunosuppressive; cardiac;
 KW nototropic; neuroprotective; gene therapy; immune disorder; wound healing;
 KW hyperproliferative disorder; renal disorder; cardiovascular disorder;
 KW respiratory disorder; neurological disease; endocrine disorder;
 KW reproductive system disorder; infectious disease;
 KW gastrointestinal; disorder.

XX Homo sapiens.

XX W0200179444-A2.

XX 25-OCT-2001.

PF 12-APR-2001; 2001WO-US12013.

XX 12-APR-2000; 2000US-229358P.

PR 25-APR-2000; 2000US-199384P.

PR 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI; 2001-616755/7i.

XX N-PSDB; ABA03057.

XX Albumin fusion proteins comprising a therapeutic protein and albumin,
 PT useful in the treating immune system disorders (e.g. transplant
 PT rejection), blood related disorders (e.g. myocardial infarction) and
 PT hyperproliferative disorders -

XX

PS Claim 1; Fig 15; 606pp; English.

XX The present invention relates to albumin fusion proteins, which comprise
 CC a therapeutic protein and albumin. The present sequence is the protein
 CC sequence for mature human serum albumin (HA), which was used to generate
 CC the fusion proteins of the present invention. The albumin fusion proteins
 CC are useful in the treatment, prevention, diagnosis, and/or detection of
 CC diseases/disorders such as immune system disorders (e.g. transplant
 CC rejection), blood related disorders (e.g. myocardial infarction),
 CC hyperproliferative disorders (e.g. childhood acute myeloid leukemia),
 CC renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
 CC arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
 CC neurological diseases (e.g. Alzheimer's disease), endocrine disorders
 CC (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),
 CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
 CC irritable bowel syndrome) and wound healing.

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 22; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKOLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
 Db 1 DAHKSEVAHRFKOLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKCEPERNECFLOHKDDNPNLRLVRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKCEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVNCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 Db 121 DVNCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCASLOKGEGERAFKAWAVARLSORPPKAEFAEVSJKLVTDLT 240
 Db 181 KLDELDEGKASSAKORLKCASLOKGEGERAFKAWAVARLSORPPKAEFAEVSJKLVTDLT 240
 QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPA 300
 Db 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPA 300
 QY 301 DPLSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYISVLLRLRAKTYETTTLEKC 360
 Db 301 DPLSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYISVLLRLRAKTYETTTLEKC 360
 QY 361 CAADDPHECYAKVDFSEFKPLVEEPONLIKONCELFQOLGEYKFORALLVRYTKVPQVST 420
 Db 361 CAADDPHECYAKVDFSEFKPLVEEPONLIKONCELFQOLGEYKFORALLVRYTKVPQVST 420
 QY 421 PTLVEVSRLGKVGSKCKHPKAEKRYPCAEEDYLSVVLNQLCVLHEKTPVSDRVTXCCTES 480
 Db 421 PTLVEVSRLGKVGSKCKHPKAEKRYPCAEEDYLSVVLNQLCVLHEKTPVSDRVTXCCTES 480
 QY 481 LVNRRPCFSALEVDVETVVPKFEFNAETTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540
 Db 481 LVNRRPCFSALEVDVETVVPKFEFNAETTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540
 QY 541 KEOLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
 Db 541 KEOLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585

RESULT 11

AAE13129
 ID AAE13129 standard; Protein; 585 AA.

XX AAE13129;

XX 28-JAN-2002 (first entry)

DE Human albumin (HA).

XX Human; albumin; HA; fusion protein; therapeutic protein; vulnereary;
 KW immune system disorder; transplant rejection; blood related disorder;
 KW myocardial infarction; hyperproliferative disorder; glomerulonephritis;
 KW childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;
 KW respiratory disorder; gene therapy; non-allergic rhinitis; noctropic;
 KW neurological disease; Alzheimer's disease; reproductive system disorder;
 KW endocrine disorder; pheocytocoma; infectious disease; antiarthritic;
 KW measles; gastrointestinal disorder; irritable bowel syndrome; syphilis;
 KW wound healing; antiinflammatory; immunosuppressive; neuroprotective;
 KW cardiant; cytostatic; antileukaemic; antirheumatic; antimicrobial;
 KW renal disorder.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Domain 54..61 /label= Loop_I
 FT Domain 76..89 /label= Loop_II
 FT Domain 92..100 /label= Loop_III
 FT Domain 170..176 /label= Loop_IV
 FT Domain 247..252 /label= Loop_V
 FT Domain 266..277 /label= Loop_VI
 FT Domain 280..288 /label= Loop_VII
 FT Domain 362..368 /label= Loop_VIII
 FT Domain 439..447 /label= Loop_IX
 FT Domain 461..475 /label= Loop_X
 FT Domain 478..486 /label= Loop_XI
 FT Domain 560..566 /label= Loop_XII
 XX
 PN WO200179443-A2.
 XX
 XX 25-OCT-2001.
 PF 12-APR-2001; 2001WO-US11924.
 XX
 XX 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 XX
 PA (HUYA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CR, Haseltine WA;
 XX
 DR WPI; 2001-616754/71.
 DR N-PSDB; AAD21638.
 XX
 DR
 XX
 PT Albumin fusion proteins comprising a therapeutic protein and albumin,
 PT useful in the treating immune system disorders (e.g. transplant
 PT rejection), blood related disorders (e.g. myocardial infarction) and
 PT hyperproliferative disorders -
 XX
 XX Claim 1; Fig 9; 380pp; English.
 PS
 CC The invention relates to albumin fusion proteins comprising therapeutic
 CC protein and human albumin (HA). Therapeutic protein fused to albumin
 CC have an extended shelf-life. The albumin fusion proteins are useful in
 CC the treatment, prevention, diagnosis and/or detection of diseases,
 CC disorders such as immune system disorders (e.g. transplant rejection),
 CC blood related disorders (e.g. myocardial infarction), hyperproliferative
 CC disorders (e.g. childhood acute myeloid leukaemia), renal disorders
 CC (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),

CC respiratory disorders (e.g. non-allergic rhinitis), neurological
 CC diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.
 CC pheocytocoma), reproductive system disorders (e.g. syphilis),
 CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
 CC irritable bowel syndrome) and wound healing. Nucleic acids encoding
 CC albumin fusion protein is used in gene therapy. The present sequence
 CC is human albumin (HA) protein.
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DS 22; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DAHKEVAHRFKDGLGEENFKALVLI AFAQYLOQCFFEDHVK-VNEVTEFAKTCVADESAAE 60
 Db 1 DAHKEVAHRFKDGLGEENFKALVLI AFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
 Qy 61 NCDKSLHT-FGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNPRLVVRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNPRLVVRPEV 120
 Qy 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
 Db 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
 Qy 181 KJDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRFPAKFAEVSCLVTDLT 240
 Db 181 KJDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRFPAKFAEVSCLVTDLT 240
 Qy 241 VHTCCGHGLLECADDRADLAKYICENODSISSKLKECCEKPLEKSHCIAEVNDMPA 300
 Db 241 VHTCCGHGLLECADDRADLAKYICENODSISSKLKECCEKPLEKSHCIAEVNDMPA 300
 Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLR-LAKTYETTLK 360
 Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLR-LAKTYETTLK 360
 Qy 361 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
 Db 361 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
 Qy 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 Qy 481 LVNRRPFSALAEVDETYVPKFNAAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 540
 Db 481 LVNRRPFSALAEVDETYVPKFNAAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 540
 Qy 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKLVVAASQAALGL 585
 Db 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKLVVAASQAALGL 585
 RESULT 12
 AAEL2403
 ID AAEL2403 standard; Protein; 585 AA.
 XX
 XX AAEL2403;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human albumin (HA).
 XX
 KW Human; albumin; HA; immune system disorder; transplant rejection;
 KW blood related disorder; myocardial infarction; glomerulonephritis;
 KW hyperproliferative disorder; childhood acute myeloid leukaemia;
 KW renal cell carcinoma; cardiovascular disorder; vulnereary; melanoma;
 KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukaemic;
 KW neurological disease; Alzheimer's disease; endocrine disorder; syphilis;
 KW pheocytocoma; reproductive system disorder; neuroprotective; syphilis;
 KW infectious disease; gastrointestinal disorder; wound healing; noctropic;

KW Irritable bowel syndrome; HIV; human immunodeficiency virus infection;
 KW cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiant;
 KW antiarthritic; antirheumatic; renal disorder; antimicrobial.

XX Homo sapiens.

XX Key Location/Qualifiers

PH Domain 54..61
 FT /label= Loop_I
 FT Domain 76..89
 FT /label= Loop_II
 FT Domain 92..100
 FT /label= Loop_III
 FT Domain 176..176
 FT /label= Loop_IV
 FT Domain 247..252
 FT /label= Loop_V
 FT Domain 266..277
 FT /label= Loop_VI
 FT Domain 280..288
 FT /label= Loop_VII
 FT Domain 362..368
 FT /label= Loop_VIII
 FT Domain 439..447
 FT /label= Loop_IX
 FT Domain 461..475
 FT /label= Loop_X
 FT Domain 478..486
 FT /label= Loop_XI
 FT Domain 560..566
 FT /label= Loop_XII

XX WO200179480-A1.

XX 25-OCT-2001.

XX 12-APR-2001; 2001WO-US11991.

XX 12-APR-2000; 2000US-229358P.

XX 25-APR-2000; 2000US-199384P.

XX 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI; 2001-616756/71.

XX N-PSDB; AAD20005.

XX Albumin fusion proteins comprising a therapeutic protein and albumin,
 PT useful in the treating metastatic renal cell carcinoma, metastatic
 PT melanoma, malignant melanoma, renal cell carcinoma, HIV (human
 PT immunodeficiency virus) or infection -

XX Claim 1; Fig 9; 394pp; English.

XX The invention relates to human albumin (HA) fusion proteins and their
 CC corresponding nucleic acid sequences. Therapeutic proteins fused to
 CC albumin or its fragments have an extended shelf-life. The albumin
 CC fusion proteins are useful in the treatment, prevention, diagnosis,
 CC and/or detection of diseases, disorders such as immune system
 CC disorders (e.g. transplant rejection), blood related disorders (e.g.
 CC myocardial infarction), hyperproliferative disorders (e.g. childhood
 CC acute myeloid leukaemia, metastatic renal cell carcinoma, metastatic
 CC melanoma, malignant melanoma, renal cell carcinoma), renal disorders
 CC (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
 CC respiratory disorders (e.g. non-allergic rhinitis), neurological
 CC diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.
 CC pheochromocytoma), reproductive system disorders (e.g. syphilis),
 CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
 CC irritable bowel syndrome), HIV (human immunodeficiency virus) infection
 CC and wound healing. Nucleic acids encoding albumin fusion protein is
 CC used in gene therapy. The present sequence is human albumin.

XX Sequence 585 AA;
 SQ Query Match 100.0%; Score 3103; DB 22; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRPKOLGEENFKALVLIIFAQYLOQCPEDHVKLVNVEVTERAKTCVADESAAE 60
 DB 1 DAHSEVAHRPKOLGEENFKALVLIIFAQYLOQCPEDHVKLVNVEVTERAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRLVRPEV 120
 QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCOAAADKAACLLP 180
 DB 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCOAAADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLLEKC 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLLEKC 360
 QY 361 CAAADPHECYAKVFEDEFKPLVEEPQNLIKONCELFEOQGEYKFONALLVRYTKKVPQVST 420
 DB 361 CAAADPHECYAKVFEDEFKPLVEEPQNLIKONCELFEOQGEYKFONALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRLGKVGSKCKGKHPKAEKMPKAEYLSVWLNQLCVLHEKTPVSDRVTKCCCTES 480
 DB 421 PTLVEVSRLGKVGSKCKGKHPKAEKMPKAEYLSVWLNQLCVLHEKTPVSDRVTKCCCTES 480
 QY 481 LVNRRPCFSALEVDTYVPKFEFNAETTFHADICTLSEKERQIKKOTALVELVKGKPKAT 540
 DB 481 LVNRRPCFSALEVDTYVPKFEFNAETTFHADICTLSEKERQIKKOTALVELVKGKPKAT 540
 QY 541 KEOLKAVNDDFAAFVKEKCKKADDDKTCFAEEGKKLVAASQAALGL 585
 DB 541 KEOLKAVNDDFAAFVKEKCKKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 13

AAE08578

ID AAE08578 standard; Protein; 585 AA.

XX

AC AAE08578;

XX

DT 19-NOV-2001 (first entry)

XX

DE Human serum albumin (HSA).

XX

KW Human; albumin; cancer; cell proliferation; drug screening; biopsy.

XX

OS Homo sapiens.

XX

PN US6274305-B1.

XX

PD 14-AUG-2001.

XX

PF 19-DEC-1996; 96US-0769746.

XX

PR 19-DEC-1996; 96US-0769746.

XX

PA (TUFT) UNIV TUFTS.

XX

PI Sonnenschein C, Soto AM;

```
XX DR NPI; 2001-540371/60.
XX DR N-PSDB; AAD11486.
XX PT Measuring human cell proliferation, useful in drug screening to
XX PT determine the potential for inhibiting cancer cell proliferation and
XX PT for evaluating biopsied tumors, comprises employing albumin-derived
XX PT peptide -
XX PS
XX PS Claim 1; Fig 1; 20pp; English.
XX CC The invention related to a method for testing cancer cells. The method is
XX CC useful for measuring human cancer cell proliferation, particularly for
XX CC determining the potential for inhibiting cancer cell's proliferation using
XX CC albumin-derived peptides. The invention is also useful for drug screening
XX CC assays, as well as for evaluating biopsied tumors. The present sequence
XX CC is human serum albumin (HSA) related to the invention.
XX SQ Sequence 585 AA;
XX
XX Query Match 100.0%; Score 3103; DB 22; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-255;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 DAHSEVAHRFKDGLGEENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHRFKDGLGEENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPFYFVAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKYLVEIARRHPFYFVAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELREDEGRKASSAKORLKCASLQKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
DB 181 KLDELREDEGRKASSAKORLKCASLQKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPJJLEKSHCIAEVENDENMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPJJLEKSHCIAEVENDENMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLRLRAKTYETTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLRLRAKTYETTLEKC 360
QY 361 CAAADPHECYAKVDFEKPPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAAADPHECYAKVDFEKPPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLGKVGSKCKKPEAKRMPCABDYLSVNLQNCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRLGKVGSKCKKPEAKRMPCABDYLSVNLQNCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVVLVKKPKAT 540
DB 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVVLVKKPKAT 540
QY 541 KEQLKAVNDQFAFVCKKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVNDQFAFVCKKADDKETCFABEGKKLVAASQAALGL 585
XX
XX RESULT 14
XX ID ABG71291
XX AC ABG71291 standard; Protein; 585 AA.
XX AC ABG71291;
XX DT 08-JAN-2003 (first entry)
XX
```

```
DE Glycosylated protein determination associated protein.
XX
XX Standard substance; accuracy control substance; glycosylated protein;
XX glycosylated albumin; fructosamine; diabetes; antidiabetic.
XX
XX Unidentified.
XX
XX JP2002243731-A.
XX
XX 28-AUG-2002.
XX
XX 21-FEB-2001; 2001JP-0045085.
XX
XX 21-FEB-2001; 2001JP-0045085.
XX
XX (KOKU-) KOKUSAI SHIYAKU KK.
XX
XX (YOSH) YOSHITOMI PHARM IND KK.
XX
XX WPI; 2002-744850/81.
XX
XX A standard substance for determination of glycosylated protein
XX including glycosylated albumin and fructosamine, used in diagnosis of
XX diabetes.
XX
XX Disclosure; Page 4; 6pp; Japanese.
XX
XX The present invention relates to a new standard and accuracy control
XX substance for determination of glycosylated protein. The invention is
XX useful for determination of glycosylated protein in the diagnosis of
XX diabetes. Glycosylated albumin and fructosamine provide favorable
XX dilution linearity. The present amino acid sequence represents the
XX glycosylated protein determination associated protein as described in
XX the invention.
XX
XX Sequence 585 AA;
XX
XX Query Match 100.0%; Score 3103; DB 23; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-255;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 DAHSEVAHRFKDGLGEENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHRFKDGLGEENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPFYFVAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKYLVEIARRHPFYFVAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELREDEGRKASSAKORLKCASLQKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
DB 181 KLDELREDEGRKASSAKORLKCASLQKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPJJLEKSHCIAEVENDENMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPJJLEKSHCIAEVENDENMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLRLRAKTYETTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLRLRAKTYETTLEKC 360
QY 361 CAAADPHECYAKVDFEKPPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAAADPHECYAKVDFEKPPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLGKVGSKCKKPEAKRMPCABDYLSVNLQNCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRLGKVGSKCKKPEAKRMPCABDYLSVNLQNCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVVLVKKPKAT 540
```

```

Db      481 LVNRRPCFSALEVDVTVPKFNAETFTFHADICTLSEKERQKKQTALVELVGHKPKAT 540
Qy      541 KEQLKAVMDDFAAVEKCCCKADDDKTCFAEGKKLVAASQAALGL 585
Db      541 KEQLKAVMDDFAAVEKCCCKADDDKTCFAEGKKLVAASQAALGL 585

RESULT 15
ABG63321
ID      ABG63321 standard; protein; 585 AA.
AC      ABG63321;
DT      27-AUG-2002 (first entry)
XX      Human serum albumin (HSA) protein.
DE      Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW      human serum albumin; HSA; cancer; reproductive disorder;
KW      digestive disorder; immune disorder; endocrine disorder;
KW      haematopoietic disorder; neural disorder; connective disorder;
KW      cytosstatic; antifertility; antiinflammatory; antidiabetic;
KW      immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
KW      neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW      osteopathic; antiarthritic.
XX      Homo sapiens.
OS      Homo sapiens.
XX      WO200177137-A1.
XX      18-OCT-2001.
XX      12-APR-2001; 2001WO-US11988.
XX      12-APR-2000; 2000US-229358P.
XX      25-APR-2000; 2000US-199384P.
XX      21-DEC-2000; 2000US-256931P.
XX      (HUMA-) HUMAN GENOME SCI INC.
XX      Rosen CA, Haseltine WA;
PI      WPI; 2002-010886/01.
XX      N-PSDB; ABK93280.
XX      New fusion protein for treating disease e.g. diabetes comprises an
PT      albumin fused to a therapeutic protein -
XX      Claim 1; Fig 15; 2102pp; English.
XX      The present invention relates to albumin fusion proteins comprising a
XX      therapeutic protein X and human albumin (HA, also known as human serum
XX      albumin, HSA). The proteins are useful for treating a disease or
XX      disorder that may be modulated by therapeutic protein X. The albumin
XX      extends the shelf-life of protein X, and may increase its biological
XX      in vitro/in vivo activity. The protein is useful for treating and
XX      diagnosing disorders such as cancer, reproductive disorders, digestive
XX      disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
XX      (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
XX      (e.g. diabetes), haematopoietic disorders, neural disorders
XX      (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
XX      encephalomyelitis, meningitis, schizophrenia), and connective disorders
XX      (e.g. osteoporosis, arthritis). The present sequence represents HSA
XX      (HA) protein.
XX      Sequence 585 AA;
SQ      Query Match 100.0%; Score 3103; DB 23; Length 585;
XX      Best Local Similarity 100.0%; Pred. No. 9.5e-255;
XX      Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      1 DAHKSEVAHREFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

```

Search completed: October 27, 2003, 15:04:45

Job time : 89 secs

```

Db      1 DAHKSEVAHREFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Qy      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGEPERNECFLOHKODNPNLPLVREPV 120
Db      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGEPERNECFLOHKODNPNLPLVREPV 120
Qy      121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQQAADKAACLLP 180
Db      121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQQAADKAACLLP 180
Qy      181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
Db      181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
Qy      241 VHTCCHGDLLECADDDRADLAKYICENODSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
Db      241 VHTCCHGDLLECADDDRADLAKYICENODSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
Qy      301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYVSVLLLRLAITYETTTLEKC 360
Db      301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYVSVLLLRLAITYETTTLEKC 360
Qy      361 CAAADPHECYAKVDFEPKPLVBEFPONLIKONCELLFEQLGEYKFONALLVRYTKVPQVST 420
Db      361 CAAADPHECYAKVDFEPKPLVBEFPONLIKONCELLFEQLGEYKFONALLVRYTKVPQVST 420
Qy      421 PTLVEVSRNLGKVGSKCKCKHPKEMPCAEDYLSVLNLQCLVHLHEKTPVSDRVTKCCTES 480
Db      421 PTLVEVSRNLGKVGSKCKCKHPKEMPCAEDYLSVLNLQCLVHLHEKTPVSDRVTKCCTES 480
Qy      481 LVNRRPCFSALEVDVTVPKFNAETFTFHADICTLSEKERQKKQTALVELVGHKPKAT 540
Db      481 LVNRRPCFSALEVDVTVPKFNAETFTFHADICTLSEKERQKKQTALVELVGHKPKAT 540
Qy      541 KEQLKAVMDDFAAVEKCCCKADDDKTCFAEGKKLVAASQAALGL 585
Db      541 KEQLKAVMDDFAAVEKCCCKADDDKTCFAEGKKLVAASQAALGL 585

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:03:17 ; Search time 29 Seconds
(without alignments)
853.512 Million cell updates/sec

Title: US-09-832-501-18

Perfect score: 3103
Sequence: 1 DAHSEVAHRFKDLGEENFK.....TCFABEGKKLVAAASQAALG 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	1 US-08-153-799-14	Sequence 14, Appli
2	3103	100.0	585	2 US-08-702-572-2	Sequence 2, Appli
3	3103	100.0	585	3 US-08-769-746-2	Sequence 2, Appli
4	3103	100.0	610	4 US-08-797-689-2	Sequence 2, Appli
5	3103	100.0	783	5 US-08-256-938-2	Sequence 2, Appli
6	3103	100.0	787	6 US-08-256-938-4	Sequence 4, Appli
7	3103	100.0	787	7 US-08-797-689-16	Sequence 16, Appli
8	3099	99.9	609	1 US-08-222-619-3	Sequence 3, Appli
9	3099	99.9	609	1 US-08-433-037-4	Sequence 4, Appli
10	3099	99.9	609	4 US-08-897-956A-2	Sequence 2, Appli
11	3095	99.9	609	5 PCT-US95-04075-3	Sequence 3, Appli
12	3095	99.7	978	4 US-08-897-956A-3	Sequence 3, Appli
13	3093	99.7	585	2 US-08-448-196A-3	Sequence 3, Appli
14	3093	99.7	585	2 US-08-984-176-1	Sequence 5, Appli
15	2458.5	79.2	583	1 US-08-448-196A-5	Sequence 5, Appli
16	2450.5	79.0	583	1 US-08-448-196A-6	Sequence 4, Appli
17	2432.5	78.4	583	1 US-08-448-196A-7	Sequence 6, Appli
18	2426	78.2	584	1 US-08-448-196A-8	Sequence 7, Appli
19	2389	77.0	582	1 US-08-134-638-1	Sequence 1, Appli
20	1249.5	40.3	609	1 US-08-222-619-4	Sequence 4, Appli
21	1249.5	40.3	609	5 PCT-US95-04075-4	Sequence 4, Appli
22	1206.5	38.9	590	2 US-08-377-309-2	Sequence 2, Appli
23	1206.5	38.9	590	3 US-09-186-723-2	Sequence 2, Appli
24	1206.5	38.9	590	4 US-08-505-012-5	Sequence 5, Appli
25	1206.5	38.9	590	4 US-09-186-949A-3	Sequence 3, Appli
26	1206.5	38.9	590	4 US-08-758-757-2	Sequence 2, Appli
27	1206.5	38.9	590	5 PCT-US96-00996-5	Sequence 5, Appli

28 1206.5 38.9 609 4 US-09-186-949A-2 Sequence 2, Appli
29 1164.5 37.5 579 1 US-08-448-196A-8 Sequence 8, Appli
30 1055 34.0 599 1 US-08-222-619-2 Sequence 2, Appli
31 1055 34.0 599 3 US-08-221-767-24 Sequence 24, Appli
32 1055 34.0 599 5 PCT-US95-04075-2 Sequence 2, Appli
33 926 29.8 393 2 US-08-377-309-7 Sequence 7, Appli
34 926 29.8 393 3 US-09-186-723-7 Sequence 7, Appli
35 926 29.8 393 4 US-08-505-012-10 Sequence 10, Appli
36 926 29.8 393 4 US-09-186-949A-8 Sequence 8, Appli
37 926 29.8 393 4 US-08-758-757-7 Sequence 7, Appli
38 926 29.8 393 5 PCT-US96-00996-10 Sequence 10, Appli
39 777 25.0 324 4 US-08-505-012-12 Sequence 12, Appli
40 777 25.0 324 5 PCT-US96-00996-22 Sequence 12, Appli
41 777 25.0 325 2 US-08-377-309-8 Sequence 8, Appli
42 777 25.0 325 3 US-09-186-723-8 Sequence 8, Appli
43 777 25.0 325 4 US-08-505-012-11 Sequence 11, Appli
44 777 25.0 325 4 US-09-186-949A-9 Sequence 9, Appli
45 777 25.0 325 4 US-08-758-757-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-06-153-799-14
; Sequence 14, Application US/08153793
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 655 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Region
; LOCATION: 369..419
; OTHER INFORMATION: /note="Alternative C-termini of
; OTHER INFORMATION: HSA(1-n)"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..585
; OTHER INFORMATION: /note="Amino acid sequence of
; OTHER INFORMATION: natural HSA"
;
US-08-153-799-14

```

```

Query Match      100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOQPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOQPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQADKAAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQADKAAACLLP 180
QY 181 KLDELRDGKASSAKQRLKASLQKFGERAFAKAWAVARLSORFFKAPFAEVSKLVTDLTK 240
DB 181 KLDELRDGKASSAKQRLKASLQKFGERAFAKAWAVARLSORFFKAPFAEVSKLVTDLTK 240
QY 241 VHTCCHGDLLECADRADLAKYICENODSISKKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECADRADLAKYICENODSISKKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDYCKNYAEAKDVFLGMFLYEAARRHPDYSVVLLRLAKTYETTLTK 360
DB 301 DLPSLAADFVESKDYCKNYAEAKDVFLGMFLYEAARRHPDYSVVLLRLAKTYETTLTK 360
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQYGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQYGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPKAMPKPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKKHPKAMPKPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETYPKPFNAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDVETYPKPFNAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
QY 541 KEQLKAVMDQDFAAFVEKCKADDDKTCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDQDFAAFVEKCKADDDKTCFAEEGKKLVAAASQAALGL 585

```

```

RESULT 2
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue

```

```

; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CR0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-702-572-2

```

```

Query Match      100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOQPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOQPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQADKAAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQADKAAACLLP 180
QY 181 KLDELRDGKASSAKQRLKASLQKFGERAFAKAWAVARLSORFFKAPFAEVSKLVTDLTK 240
DB 181 KLDELRDGKASSAKQRLKASLQKFGERAFAKAWAVARLSORFFKAPFAEVSKLVTDLTK 240
QY 241 VHTCCHGDLLECADRADLAKYICENODSISKKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECADRADLAKYICENODSISKKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDYCKNYAEAKDVFLGMFLYEAARRHPDYSVVLLRLAKTYETTLTK 360
DB 301 DLPSLAADFVESKDYCKNYAEAKDVFLGMFLYEAARRHPDYSVVLLRLAKTYETTLTK 360
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQYGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQYGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPKAMPKPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKKHPKAMPKPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETYPKPFNAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDVETYPKPFNAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540

```

QY 541 KEOLKAVMDFFAFAVEKCKKADDDKTCFAEKGKLVAAQAALGL 585
 DB 541 KEOLKAVMDFFAFAVEKCKKADDDKTCFAEKGKLVAAQAALGL 585
 RESULT 3
 US-08-769-746-2
 ; Sequence 2, Application US/08769746
 ; Patent No. 6274305
 ; GENERAL INFORMATION:
 ; APPLICANT: Sonnenschein, Carlos
 ; APPLICANT: Soto, Ana M.
 ; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Medlen & Carroll, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent-In Release #1.0, Version #1.10
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/769,746
 ; FILING DATE: 19-DEC-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Carroll, Peter G.
 ; REGISTRATION NUMBER: 32,837
 ; REFERENCE/DOCKET NUMBER: MBRI-02584
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 585 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-769-746-2

Query Match 100.0%; Score 3103; DB 3; Length 585;
 Best Local Similarity 100.0%; Pred. NO. 9.3e-287;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAKSEVAHRFKDLGKLVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
 DB 1 DAKSEVAHRFKDLGKLVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKKDDNPNLRLVRPEV 120
 QY 121 DVMTCTADHNEETELKYLVEIAERHPYFYAPELLFFAKRYKAFTCCOAAKACLLP 180
 DB 121 DVMTCTADHNEETELKYLVEIAERHPYFYAPELLFFAKRYKAFTCCOAAKACLLP 180
 QY 181 KLDELREGKASSAKQRLKCSAQKGERAFKAWAVARLSQRPKAEFAEVSKLVDLTAK 240
 DB 181 KLDELREGKASSAKQRLKCSAQKGERAFKAWAVARLSQRPKAEFAEVSKLVDLTAK 240
 QY 241 VHTCCGGDLLEADRADLAKYTCENQDSISSKLECKECPKLEKSHCHIAEVNDEMPA 300
 DB 241 VHTCCGGDLLEADRADLAKYTCENQDSISSKLECKECPKLEKSHCHIAEVNDEMPA 300
 QY 301 DLPSLAADFVESKVCQYAEAKDVFGLMELYEARHPDYSVVLLRLAKTYETTTLEKC 360
 DB 301 DLPSLAADFVESKVCQYAEAKDVFGLMELYEARHPDYSVVLLRLAKTYETTTLEKC 360

QY 361 CAADPHCEYAKVDFEKLPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADPHCEYAKVDFEKLPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRLNGLKVGSKCKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRLNGLKVGSKCKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGKPKAT 540
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGKPKAT 540
 QY 541 KEOLKAVMDFFAFAVEKCKKADDDKTCFAEKGKLVAAQAALGL 585
 DB 541 KEOLKAVMDFFAFAVEKCKKADDDKTCFAEKGKLVAAQAALGL 585
 RESULT 4
 US-08-797-689-2
 ; Sequence 2, Application US/08797689
 ; Patent No. 5876969
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleer, Reinhard
 ; APPLICANT: Fournier, Alain
 ; APPLICANT: Guitton, Jean-Dominique
 ; APPLICANT: Jung, Gerard
 ; APPLICANT: Yeh, Patrice
 ; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 ; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3C43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: System 7.1
 ; SOFTWARE: Word 5.1 (patentin)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/797,689
 ; FILING DATE: 31-JAN-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/256,927
 ; FILING DATE: 28-JUL-1994
 ; APPLICATION NUMBER: FR 92/01064
 ; FILING DATE: 31-JAN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/FR93/00085
 ; FILING DATE: 28-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith Ph.D., Julie K.
 ; REGISTRATION NUMBER: P-38,619
 ; REFERENCE/DOCKET NUMBER: ST92006-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610) 454-3839
 ; TELEFAX: (610) 454-3808
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 610 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-797-689-2

Query Match 100.0%; Score 3103; DB 2; Length 610;
 Best Local Similarity 100.0%; Pred. NO. 9.8e-287;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 DAHSEVAHRFKDLEENFKALVLIIFAQYLQOCPEFHVHVLKNEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLEENFKALVLIIFAQYLQOCPEFHVHVLKNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 144
QY 121 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELIRDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPKAEFAEVSCLVTDLT 240
Db 205 KLDELIRDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAEARRHPDYSVVLRLAKTYETTLEK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAEARRHPDYSVVLRLAKTYETTLEK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
Db 385 CAADPHCEYAKVDFEFPKPLVEEPONLIKQNCLEFQOLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTPKFNAEFTTFHADICTLSEKEROIKKQATALVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDYTPKFNAEFTTFHADICTLSEKEROIKKQATALVELVHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDETCFAEBGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDDFAAFVEKCKCKADDETCFAEBGKKLVAASQAALGL 609

```

RESULT 5

```

US-08-256-938-2
; Sequence 2, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne

```

```

; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: ST92007-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-938-2

```

```

Query Match 100.0%; Score 3103; DB 1; Length 783;
Best Local Similarity 100.0%; Pred No. 1,4e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLEENFKALVLIIFAQYLQOCPEFHVHVLKNEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLEENFKALVLIIFAQYLQOCPEFHVHVLKNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 144
QY 121 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELIRDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPKAEFAEVSCLVTDLT 240
Db 205 KLDELIRDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAEARRHPDYSVVLRLAKTYETTLEK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAEARRHPDYSVVLRLAKTYETTLEK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
Db 385 CAADPHCEYAKVDFEFPKPLVEEPONLIKQNCLEFQOLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTPKFNAEFTTFHADICTLSEKEROIKKQATALVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDYTPKFNAEFTTFHADICTLSEKEROIKKQATALVELVHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDETCFAEBGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDDFAAFVEKCKCKADDETCFAEBGKKLVAASQAALGL 609

```

RESULT 6

```

US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA

```

COUNTRY: USA
 ZIP: 19426
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.0 (PatentIn)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/256,938
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 92/01065
 FILING DATE: 31-JAN-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Goodman, Rosanne
 REGISTRATION NUMBER: 32,534
 REFERENCE/DOCKET NUMBER: ST92007-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3817
 TELEFAX: (610) 454-3808
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 787 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-256-938-4

Query Match 100.0%; Score 3103; DB 1; Length 787;
 Best Local Similarity 100.0%; Pred. No. 1.4e-286;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAKSEVAHREKDLGKLVIAFAQYLOQCPEFHVKLVNEVTEFAKTCVADES	60
DB	203	DAKSEVAHREKDLGKLVIAFAQYLOQCPEFHVKLVNEVTEFAKTCVADES	262
QY	61	NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	120
DB	263	NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	322
QY	121	DVMCTAFHNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAAACLLP	180
DB	323	DVMCTAFHNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAAACLLP	382
QY	181	KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSORFPKAEFAEVS	240
DB	383	KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSORFPKAEFAEVS	442
QY	241	VHTECHGDLLECAADRADLAKYICENQDISSKLKECCEKPLLEKSHCIAEVENDEMPA	300
DB	443	VHTECHGDLLECAADRADLAKYICENQDISSKLKECCEKPLLEKSHCIAEVENDEMPA	502
QY	301	DLPSLAADPVESKDVCKNTAEAKDVLGMLFVEYARRHPDYSVVLRLRAKTYETTLK	360
DB	503	DLPSLAADPVESKDVCKNTAEAKDVLGMLFVEYARRHPDYSVVLRLRAKTYETTLK	562
QY	361	CAAADPHECYAKVDFEFKPLVSEPNLIIKONCELFQGEYKFNALLVRYTKYPOVST	420
DB	563	CAAADPHECYAKVDFEFKPLVSEPNLIIKONCELFQGEYKFNALLVRYTKYPOVST	622
QY	421	PTLVEYSRNLGKVGSKCKHPBAKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES	480
DB	623	PTLVEYSRNLGKVGSKCKHPBAKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES	682
QY	481	LNNRRCFSALEVDYTPKFNASTFTTHADICTLSEKEROIKKQATALVELVGHKPKAT	540
DB	683	LNNRRCFSALEVDYTPKFNASTFTTHADICTLSEKEROIKKQATALVELVGHKPKAT	742
QY	541	KEQLKAVMDQFAAFVEKCKKADDKETCFPAEEGKLVAAASQAALG	585
DB	743	KEQLKAVMDQFAAFVEKCKKADDKETCFPAEEGKLVAAASQAALG	787

RESULT 7
 US-08-797-689-16
 ; Sequence 16, Application US/08797689
 ; Patent No. 5876969
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleer, Reinhard
 ; APPLICANT: Fournier, Alain
 ; APPLICANT: Guitton, Jean-Dominique
 ; APPLICANT: Jung, Gerard
 ; APPLICANT: Yeh, Patrice
 ; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 ; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 ; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3C43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: System 7.1
 ; SOFTWARE: Word 5.1 (PatentIn)
 ; CURRENT APPLICATION DATA: US/08/797,689
 ; APPLICATION NUMBER: US/08/797,689
 ; FILING DATE: 31-JAN-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/256,927
 ; FILING DATE: 28-JUL-1994
 ; APPLICATION NUMBER: FR 92/01064
 ; FILING DATE: 31-JAN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/FR93/00085
 ; FILING DATE: 28-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith Ph.D., Julie K.
 ; REGISTRATION NUMBER: P-38,619
 ; REFERENCE/DOCKET NUMBER: ST92006-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610) 454-3839
 ; TELEFAX: (610) 454-3808
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 787 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-797-689-16

Query Match 100.0%; Score 3103; DB 2; Length 787;
 Best Local Similarity 100.0%; Pred. No. 1.4e-286;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHSEVAHREKDLGKLVIAFAQYLOQCPEFHVKLVNEVTEFAKTCVADES	60
DB	203	DAHSEVAHREKDLGKLVIAFAQYLOQCPEFHVKLVNEVTEFAKTCVADES	262
QY	61	NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	120
DB	263	NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	322
QY	121	DVMCTAFHNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAAACLLP	180
DB	323	DVMCTAFHNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAAACLLP	382
QY	181	KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSORFPKAEFAEVS	240
DB	383	KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSORFPKAEFAEVS	442

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPEP-LEKSHCIAEVENDEMPA 300
DB 443 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPEP-LEKSHCIAEVENDEMPA 502
QY 301 DLPSLAADFVESKDVCKVKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTTLEK 360
DB 503 DLPSLAADFVESKDVCKVKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTTLEK 562
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPOVST 420
DB 563 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPOVST 622
QY 421 PTLVEVSRLNKGKVGSKCKHPEAKMPCAEEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
DB 623 PTLVEVSRLNKGKVGSKCKHPEAKMPCAEEDYLSVNLQCLVHEKTPVSDRVTKCCTES 662
QY 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVKKPKAT 540
DB 683 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVKKPKAT 742
QY 541 KEQLKAVMDFFAAVFEVKCKKADOKKETCFABEGKKLVAASQAALGL 585
DB 743 KEQLKAVMDFFAAVFEVKCKKADOKKETCFABEGKKLVAASQAALGL 787

RESULT 8

US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichensstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/222.619
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPYQLOQCFEDHVKLVNEVTEFAKTCVADESA 60
DB 25 DAHKEVAHRFKDLGEENFKALVLIAPYQLOQCFEDHVKLVNEVTEFAKTCVADESA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVPEV 120

DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVPEV 144
QY 121 DVMCTAFHDEETFLKYLVEIARRHPYFYVAPELLFFAKYKAAFTSCCOAADAACCLLP 180
DB 145 DVMCTAFHDEETFLKYLVEIARRHPYFYVAPELLFFAKYKAAFTSCCOAADAACCLLP 204
QY 181 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240
DB 205 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPEP-LEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPEP-LEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKVKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTTLEK 360
DB 325 DLPSLAADFVESKDVCKVKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTTLEK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPOVST 420
DB 385 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPOVST 444
QY 421 PTLVEVSRLNKGKVGSKCKHPEAKMPCAEEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNKGKVGSKCKHPEAKMPCAEEDYLSVNLQCLVHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVKKPKAT 540
DB 505 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVKKPKAT 564
QY 541 KEQLKAVMDFFAAVFEVKCKKADOKKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDFFAAVFEVKCKKADOKKETCFABEGKKLVAASQAALGL 609

RESULT 9
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Brierley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Tschopp, Juerg P.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; TITLE OF INVENTION: PICHIA PASTORIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 03-MAY-1995
; APPLICATION NUMBER: US/08/433,037
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9108Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-296;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSEVAHREFKDLGEBNFKALVLIAPAFYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHKSEVAHREFKDLGEBNFKALVLIAPAFYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEERNECFLOHKDDPNLPLRVREP 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEERNECFLOHKDDPNLPLRVREP 144
QY 121 DVNCTAFHNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAAACLLP 180
DB 145 DVNCTAFHNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAAACLLP 204
QY 181 KLDELDEGKASSAKORLKCSLQKFGERAFAKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
DB 205 KLDELDEGKASSAKORLKCSLQKFGERAFAKAWAVARLSQRPFAEFAEVSKLVTDLTK 264
QY 241 VHTTECHGDLLLECADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEPA 300
DB 265 VHTTECHGDLLLECADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 384
QY 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 385 CAAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHKPKAT 540
DB 505 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCKADKCTCFABEGKGLVAASQAALGL 585
DB 565 KEQLKAVMDDFAAFVEKCKCKADKCTCFABEGKGLVAASQAALGL 609

RESULT 10
US-08-897-956A-2
Sequence 2, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT APPLICATION NUMBER: US/08/897,956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 99.9%; Score 3099; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHREFKDLGEBNFKALVLIAPAFYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHKSEVAHREFKDLGEBNFKALVLIAPAFYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEERNECFLOHKDDPNLPLRVREP 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEERNECFLOHKDDPNLPLRVREP 144
QY 121 DVNCTAFHNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAAACLLP 180
DB 145 DVNCTAFHNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAAACLLP 204
QY 181 KLDELDEGKASSAKORLKCSLQKFGERAFAKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
DB 205 KLDELDEGKASSAKORLKCSLQKFGERAFAKAWAVARLSQRPFAEFAEVSKLVTDLTK 264
QY 241 VHTTECHGDLLLECADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEPA 300
DB 265 VHTTECHGDLLLECADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 384
QY 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 385 CAAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHKPKAT 540
DB 505 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCKADKCTCFABEGKGLVAASQAALGL 585
DB 565 KEQLKAVMDDFAAFVEKCKCKADKCTCFABEGKGLVAASQAALGL 609

RESULT 1:
PCT-US95-04075-3
Sequence 3, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/PRC
STREET: 1840 DeHavilland Drive
City: Thousand Oaks
STATE: California
COUNTRY: U.S.
Zip: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04075
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match 99.9%; Score 3099; DB 5; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 3;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOQPFEDHVKLVNVEVTEFAKTCVADESAS 60
DB 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOQPFEDHVKLVNVEVTEFAKTCVADESAS 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 144
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCAASQKFGGERAFKAWAVARLSQRFPAKAEFAEVSCLVTLTK 240
DB 205 KLDELDEGKASSAKQRLKCAASQKFGGERAFKAWAVARLSQRFPAKAEFAEVSCLVTLTK 264
QY 241 VHTCCHGDLLECADRADLAKYICENODS:SSKLKECEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCHGDLLECADRADLAKYICENODS:SSKLKECEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLYVEYARRHPDYVSVLLRLAKTYETTLK 360
DB 325 DLPSLAADPVESKDVCKNYAEAKDVLGMFLYVEYARRHPDYVSVLLRLAKTYETTLK 384
QY 361 CAADPHECYAKVDFEPLVEEPQNLKQNCLEPQOLGEYKFNALLVRYTKVQVST 420
DB 385 CAADPHECYAKVDFEPLVEEPQNLKQNCLEPQOLGEYKFNALLVRYTKVQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALVDETYVPKFNAEFTTFHADICTLSEKERQIKQTALVELVHKPKAT 540
DB 505 LVNRRPCFSALVDETYVPKFNAEFTTFHADICTLSEKERQIKQTALVELVHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFAEEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFAEEGKKLVAASQAALGL 609

RESULT 12
US-08-897-956A-3
Sequence 3, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT APPLICATION NUMBER: US/08/897,956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 978
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion polypeptide

US-08-897-956A-3

Query Match 99.7%; Score 3095; DB 4; Length 978;
Best Local Similarity 99.8%; Pred. No. 1.1e-285;
Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOQPFEDHVKLVNVEVTEFAKTCVADESAS 60
DB 212 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOQPFEDHVKLVNVEVTEFAKTCVADESAS 271
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 120
DB 272 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 331
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 332 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 391
QY 181 KLDELDEGKASSAKQRLKCAASQKFGGERAFKAWAVARLSQRFPAKAEFAEVSCLVTLTK 240
DB 392 KLDELDEGKASSAKQRLKCAASQKFGGERAFKAWAVARLSQRFPAKAEFAEVSCLVTLTK 451
QY 241 VHTCCHGDLLECADRADLAKYICENODS:SSKLKECEKPLLEKSHCIAEVENDEMPA 300
DB 452 VHTCCHGDLLECADRADLAKYICENODS:SSKLKECEKPLLEKSHCIAEVENDEMPA 511
QY 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLYVEYARRHPDYVSVLLRLAKTYETTLK 360
DB 512 DLPSLAADPVESKDVCKNYAEAKDVLGMFLYVEYARRHPDYVSVLLRLAKTYETTLK 571
QY 361 CAADPHECYAKVDFEPLVEEPQNLKQNCLEPQOLGEYKFNALLVRYTKVQVST 420
DB 572 CAADPHECYAKVDFEPLVEEPQNLKQNCLEPQOLGEYKFNALLVRYTKVQVST 631
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 632 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 691
QY 481 LVNRRPCFSALVDETYVPKFNAEFTTFHADICTLSEKERQIKQTALVELVHKPKAT 540
DB 692 LVNRRPCFSALVDETYVPKFNAEFTTFHADICTLSEKERQIKQTALVELVHKPKAT 751
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFAEEGKKLVAASQAALG 584
DB 752 KEQLKAVMDDDFAAFVEKCKKADDKETCFAEEGKKLVAASQAALG 795

RESULT 13
US-08-448-196A-3
Sequence 3, Application US/08448196A
Patent No. 5780594
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN
TITLE OF INVENTION: RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-3

```
Query Match          99.7%; Score 3093; DB 1; Length 585;
Best Local Similarity 99.7%; Pred. No. 8.3e-286;
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLSGEENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLSGEENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFPHDNEETFLKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMTAFPHDNEETFLKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEFAEVSCLVTDLTK 240
DB 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEFAEVSCLVTDLTK 240
QY 241 VHTCCHGDLLECCADRADLAKYICENODSISSKLKCECEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECCADRADLAKYICENODSISSKLKCECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKVCNKYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLTK 360
DB 301 DLPSLAADFVESKVCNKYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLTK 360
QY 361 CAADHPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADHPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETVVPKFEAFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
DB 481 LVNRRPCFSALEVDETVVPKFEAFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
```

RESULT 14
US-08-984-176-1
Sequence 1, Application: US/08984176
Patent No. 5948609
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C
APPLICANT: HO, JOSEPH X
APPLICANT: RUKER, FLORIAN
TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176
CURRENT APPLICATION NUMBER: US/08/984,176
CURRENT FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 585
TYPE: PRT
ORGANISM: Homo sapiens
US-08-984-176-1

```
Query Match          99.7%; Score 3093; DB 2; Length 585;
Best Local Similarity 99.7%; Pred. No. 8.3e-286;
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLSGEENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLSGEENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFPHDNEETFLKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMTAFPHDNEETFLKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEFAEVSCLVTDLTK 240
DB 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEFAEVSCLVTDLTK 240
QY 241 VHTCCHGDLLECCADRADLAKYICENODSISSKLKCECEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECCADRADLAKYICENODSISSKLKCECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKVCNKYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLTK 360
DB 301 DLPSLAADFVESKVCNKYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLTK 360
QY 361 CAADHPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADHPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETVVPKFEAFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
DB 481 LVNRRPCFSALEVDETVVPKFEAFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
```

RESULT 15
US-08-448-196A-5
Sequence 5, Application: US/08448196A
Patent No. 5780594
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
TITLE OF INVENTION: RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812

Job time : 32 secs

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-5
```

```
Query Match 79.2%; Score 2458.5; DB 1; Length 583;
Best Local Similarity 75.8%; Pred. No. 2.1e-225;
Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHSEVAHREFKDLGEENFKALVLIATFAQYLQOCFFEDHVKLVNEVTEPAKTCVADESAE 60
DB 1 DTHKEIAHRPNNDLGEKHFGLVLVAFSQYLQOCFFEDHVKLVNEVTEPAKCAADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 119

QY 121 DVMCTAFHDSNETFLKLYLRIARRHPYFAPELIFFAKRYKAAFTCCCAADKACLLP 180
DB 120 DAQCAAFQEDPDKFLGKLYEVARRHPYFYPPELLFHAEYKADFTCCPADDKLACLP 179

QY 181 KLDELDRGKASSAKQRLKASQKQFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 240
DB 180 KLDALKERILLSSAKERLKCSSFQNFGERAVKANSVARLSQKFPKADFAEVSKIIVTLTK 239

QY 241 VHTCECHGDLLECADDRADLAKYICENODSISSKLKECCEKPLEKSHCIAEYENDEMPA 300
DB 240 VHKECCHGDLLECADDRADLAKYICEHODS-SGKLKACCDKPLQKSHCIAEVKEDDLPS 299

QY 301 DLPSLAADFVESKDYCKNYBAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 300 DIPALLAADFADKEICKHYKQAKDVFLGTFLEYYSRRHPDYSVLLRLAKTYEATLEK 359

QY 361 CAADPHCEYAKVPDEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
DB 360 CAEADPPACRYTFVDFQTFPLVEEPKSLVKQNCDFEEVGEYDFQNALIVRYTKKAPQVST 419

QY 421 PTLVEVSRNLGKVGSKCKHPKPEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 420 PTLVEIGRTLKGKVGSRCKLPSEERLPCSENHLAALNRLCVLHEKTPVSEKIKKCCDES 479

QY 481 LVNRRPCFSALEDETYVPKFNASTFTFHADICTLSEKERQIKQTALVELVVKHKPAT 540
DB 480 LAERRPCFSALEDEGYVPKFEKATFTFHADICTLPEDEKQIKKQSALAEVLVVKHKPAT 539

QY 541 KEQLKAVMDDFAAFVEKCKADDKETCTFAEGKKLVAAASQAL 583
DB 540 KEQLKTVLGNFSAFVAKCCGREDKEACFAEBGPKLVASSQAL 582
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:07:18 ; Search time 73 Seconds
(without alignments)
1341.979 Million cell updates/sec

Title: US-09-832-501-18

Perfect score: 3103

Sequence: 1 DAHSEVAHRFDLGEENFK.....TCFAEKGKLVAAQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
3: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
4: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
7: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
8: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
9: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
10: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
11: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
12: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
13: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
14: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
15: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
16: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
17: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:
18: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	10	US-09-929-552-2 Sequence 2, Appli
2	3103	100.0	585	11	US-09-932-613-445 Sequence 445, App
3	3103	100.0	585	11	US-09-984-010-26 Sequence 26, Appl
4	3103	100.0	585	11	US-09-833-041-18 Sequence 18, Appl
5	3103	100.0	585	12	US-10-153-604A-5 Sequence 5, Appli
6	3103	100.0	585	12	US-09-833-117-18 Sequence 18, Appl
7	3103	100.0	585	12	US-10-319-263-1 Sequence 1, Appli
8	3103	100.0	585	12	US-10-319-263-2 Sequence 2, Appli
9	3103	100.0	585	12	US-10-414-469-1 Sequence 1, Appli
10	3103	100.0	585	12	US-10-414-469-2 Sequence 2, Appli
11	3103	100.0	585	12	US-09-932-322-445 Sequence 445, App
12	3103	100.0	585	12	US-10-413-831-1 Sequence 1, Appli
13	3103	100.0	585	12	US-10-413-831-2 Sequence 2, Appli
14	3103	100.0	585	14	US-10-153-064-5 Sequence 5, Appli
15	3103	100.0	604	11	US-09-984-010-7 Sequence 7, Appli

16	3103	100.0	609	11	US-09-919-039-370 Sequence 370, App
17	3103	100.0	609	12	US-10-153-604A-7 Sequence 7, Appli
18	3103	100.0	609	12	US-10-365-623-23 Sequence 23, Appl
19	3103	100.0	609	14	US-10-153-064-7 Sequence 7, Appli
20	3103	100.0	610	10	US-09-984-186-2 Sequence 2, Appli
21	3103	100.0	610	15	US-10-237-667-2 Sequence 2, Appli
22	3103	100.0	610	15	US-10-237-708-2 Sequence 2, Appli
23	3103	100.0	610	15	US-10-237-866-2 Sequence 2, Appli
24	3103	100.0	610	15	US-10-237-871-2 Sequence 2, Appli
25	3103	100.0	610	15	US-10-237-871-2 Sequence 2, Appli
26	3103	100.0	651	12	US-10-153-604A-133 Sequence 133, App
27	3103	100.0	651	14	US-10-153-064-133 Sequence 133, App
28	3103	100.0	652	12	US-10-153-604A-132 Sequence 132, App
29	3103	100.0	652	14	US-10-153-064-132 Sequence 132, App
30	3103	100.0	653	12	US-10-153-604A-131 Sequence 131, App
31	3103	100.0	653	14	US-10-153-064-131 Sequence 131, App
32	3103	100.0	656	12	US-10-153-604A-130 Sequence 130, App
33	3103	100.0	656	14	US-10-153-064-130 Sequence 130, App
34	3103	100.0	676	12	US-10-153-604A-127 Sequence 127, App
35	3103	100.0	676	12	US-10-153-604A-129 Sequence 129, App
36	3103	100.0	676	14	US-10-153-064-127 Sequence 127, App
37	3103	100.0	676	14	US-10-153-064-129 Sequence 129, App
38	3103	100.0	677	12	US-10-153-604A-125 Sequence 125, App
39	3103	100.0	677	14	US-10-153-064-125 Sequence 125, App
40	3103	100.0	680	12	US-10-153-604A-123 Sequence 123, App
41	3103	100.0	680	14	US-10-153-064-123 Sequence 123, App
42	3103	100.0	787	10	US-09-984-186-16 Sequence 16, Appl
43	3103	100.0	787	15	US-10-237-667-16 Sequence 16, Appl
44	3103	100.0	787	15	US-10-237-708-16 Sequence 16, Appl
45	3103	100.0	787	15	US-10-237-866-16 Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sornenschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/929,552
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIORITY INFORMATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match      100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIATFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDGLGEENFKALVLIATFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEKNECFLOHKDNDNPNLRLVPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEKNECFLOHKDNDNPNLRLVPEV 120
QY 121 DVMTAFHNDNEETFLKKYLYEIAARRHPYFYAPPELLFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHNDNEETFLKKYLYEIAARRHPYFYAPPELLFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELURDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTK 240
DB 181 KLDELURDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTK 240
QY 241 VHTTECHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTLTK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTLTK 360
QY 361 CAADPHCYAKVDFEPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
DB 361 CAADPHCYAKVDFEPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
QY 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 540
QY 541 KEQLKAVMDDFAAAFVEKCKKADDDKTCFAEBGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAAFVEKCKKADDDKTCFAEBGKGLVAASQAALGL 585

RESULT 3
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Balliance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERJM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPC)
; CURRENT APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164

```

; FILING DATE: 19-DEC-1996
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-832-501-18

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKXYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKXYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCSLQKFGGERAPKAMAVARLSQRFPKAEFAEVSKLVDLT 240
DB 181 KLDELDEGKASSAKORLKCSLQKFGGERAPKAMAVARLSQRFPKAEFAEVSKLVDLT 240
QY 241 VHTCCHGDLLECADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKVCNKYAEAKOVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKVCNKYAEAKOVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 360
QY 361 AAAADPHECYAKVDFEKLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
DB 361 AAAADPHECYAKVDFEKLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
QY 421 PTLVEVSRNLGKVGSKCKGHPKAMPKCAEDYLSVNLQCLVHLEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKGHPKAMPKCAEDYLSVNLQCLVHLEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTPKPEFNAETFTPHADICTLSEKEROIKKOTALVELVKGKPKAT 540
DB 481 LVNRRPCFSALEVDYTPKPEFNAETFTPHADICTLSEKEROIKKOTALVELVKGKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCPAEKGKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKCKADKCTCPAEKGKLVAAASQAALGL 585

RESULT 4
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF545
; CURRENT APPLICATION NUMBER: US/09/833,041
; PRIOR FILING DATE: 2001-04-12
; PRIOR FILING DATE: 2000-04-12
; PRIOR FILING DATE: 2000-04-12
; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKXYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKXYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCSLQKFGGERAPKAMAVARLSQRFPKAEFAEVSKLVDLT 240
DB 181 KLDELDEGKASSAKORLKCSLQKFGGERAPKAMAVARLSQRFPKAEFAEVSKLVDLT 240
QY 241 VHTCCHGDLLECADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKVCNKYAEAKOVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKVCNKYAEAKOVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 360
QY 361 AAAADPHECYAKVDFEKLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
DB 361 AAAADPHECYAKVDFEKLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
QY 421 PTLVEVSRNLGKVGSKCKGHPKAMPKCAEDYLSVNLQCLVHLEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKGHPKAMPKCAEDYLSVNLQCLVHLEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTPKPEFNAETFTPHADICTLSEKEROIKKOTALVELVKGKPKAT 540
DB 481 LVNRRPCFSALEVDYTPKPEFNAETFTPHADICTLSEKEROIKKOTALVELVKGKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCPAEKGKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKCKADKCTCPAEKGKLVAAASQAALGL 585

RESULT 5
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; PRIOR FILING DATE: 2002-05-24
; PRIOR FILING DATE: 2002-05-24
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens

US-10-153-604A-S

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIATAFYLOQCQPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGEENFKALVLIATAFYLOQCQPFEDHVKLVNEVTEFAKTCVADESAAE 60

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRVPRPV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRVPRPV 120

Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 180

Qy 181 KLDELDEGKASSAKORLKASQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTJK 240
Db 181 KLDELDEGKASSAKORLKASQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTJK 240

Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKKECCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKKECCEKPLLEKSHCIAEVENDEMPA 300

Qy 301 DLPSLAADFVESKVCNVAEAKDVFLGMFLVEYARRHPDYSVWLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKVCNVAEAKDVFLGMFLVEYARRHPDYSVWLLRLAKTYETTTLEKC 360

Qy 361 CAADPHECYAKVDFEFKPLVEBPONLIKONCELFEOQGEYKFNALLVRYTKVQVQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEBPONLIKONCELFEOQGEYKFNALLVRYTKVQVQVST 420

Qy 421 PTLVEVSRNLGVKSGKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGVKSGKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Qy 481 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540

Qy 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEBGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEBGKKLVAASQAALGL 585

RESULT 6

US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

RESULT 7

US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
;

US-10-319-263-1

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOQCFEDHVKLVNVEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOQCFEDHVKLVNVEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTCTAFHDNEETFLKKYLYIARHHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180
DB 121 DVMTCTAFHDNEETFLKKYLYIARHHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180
QY 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
DB 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKCECKEPPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKCECKEPPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEKLVEEPQNLIKQNCLEFQELGEYKFKQNALLVRYTKVPQVST 420
DB 361 CAADAPHECYAKVDFEKLVEEPQNLIKQNCLEFQELGEYKFKQNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLNKGVSCKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNKGVSCKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKFNAEFTTFHADICTLSEKEROIKKQOTALVELVXHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKFNAEFTTFHADICTLSEKEROIKKQOTALVELVXHKPKAT 540
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
```

RESULT 8

```
US-10-319-263-2
; Sequence 2, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
```

ORGANISM: Homo sapiens

```
FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLTATION
US-10-319-263-2
```

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOQCFEDHVKLVNVEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOQCFEDHVKLVNVEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTCTAFHDNEETFLKKYLYIARHHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180
DB 121 DVMTCTAFHDNEETFLKKYLYIARHHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180
QY 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
DB 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKCECKEPPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKCECKEPPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEKLVEEPQNLIKQNCLEFQELGEYKFKQNALLVRYTKVPQVST 420
DB 361 CAADAPHECYAKVDFEKLVEEPQNLIKQNCLEFQELGEYKFKQNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLNKGVSCKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNKGVSCKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKFNAEFTTFHADICTLSEKEROIKKQOTALVELVXHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKFNAEFTTFHADICTLSEKEROIKKQOTALVELVXHKPKAT 540
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
```

RESULT 9

```
US-10-414-469-1
; Sequence 1, Application US/10414469
; Publication No. US20030190691A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
```

```
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MOD_RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLATION
US-10-414-469-2

Query Match: 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYIQOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKEVAHRFKDLGEENFKALVLIAPAYIQOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRITYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRITYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
QY 121 DVMTAFHNDNEETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 121 DVMTAFHNDNEETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAMAVARLSQRFPAKAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAMAVARLSQRFPAKAEFAEVSKLVTDLTK 240
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300
Db 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADVESEKDVCKNYAEAKDVFLGMFLYEVARRHPDYVSVLLLLRLAKTYETTLK 360
Db 301 DLPSLAADVESEKDVCKNYAEAKDVFLGMFLYEVARRHPDYVSVLLLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFLVEEPONLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFLVEEPONLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGVGSKCCCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGVGSKCCCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQATALVELVXHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQATALVELVXHKPKAT 540
QY 541 KEOLKAVMDDDFAAFVEKCKCKADDKETCTFAEBGKCLVAASQAALGL 585
Db 541 KEOLKAVMDDDFAAFVEKCKCKADDKETCTFAEBGKCLVAASQAALGL 585

RESULT 1:
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.

Query Match: 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYIQOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKEVAHRFKDLGEENFKALVLIAPAYIQOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRITYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRITYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
QY 121 DVMTAFHNDNEETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 121 DVMTAFHNDNEETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAMAVARLSQRFPAKAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAMAVARLSQRFPAKAEFAEVSKLVTDLTK 240
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300
Db 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADVESEKDVCKNYAEAKDVFLGMFLYEVARRHPDYVSVLLLLRLAKTYETTLK 360
Db 301 DLPSLAADVESEKDVCKNYAEAKDVFLGMFLYEVARRHPDYVSVLLLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFLVEEPONLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFLVEEPONLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGVGSKCCCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGVGSKCCCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQATALVELVXHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQATALVELVXHKPKAT 540
QY 541 KEOLKAVMDDDFAAFVEKCKCKADDKETCTFAEBGKCLVAASQAALGL 585
Db 541 KEOLKAVMDDDFAAFVEKCKCKADDKETCTFAEBGKCLVAASQAALGL 585

RESULT 10
US-10-414-469-2
; Sequence 2, Application US/10414469
; Publication No. US20030190691A1
; GENERAL INFORMATION:
; APPLICANT: Bat-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905
```



```
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/832.322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-322-445

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGSENFKALVLIAPAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60
Db 1 DAHKSEVAHRFKDLGSENFKALVLIAPAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVRPEV 120
Qy 121 DVMTAFHNDNEETFLKKLYEIAHRHPFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db 121 DVMTAFHNDNEETFLKKLYEIAHRHPFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Qy 181 KLDELDRSGKASSAKQRLKCSLQKFGGERAFKAWAVARLSORFPKAEFAEVSKLVDLT 240
Db 181 KLDELDRSGKASSAKQRLKCSLQKFGGERAFKAWAVARLSORFPKAEFAEVSKLVDLT 240
Qy 241 VHTCCGHDLLCADDRAADLAKYICENQDS:SSKLKECCERPLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGHDLLCADDRAADLAKYICENQDS:SSKLKECCERPLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGFLYFYARRHPDYSVLLRLAKTYETLEK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGFLYFYARRHPDYSVLLRLAKTYETLEK 360
Qy 361 CAAADPHECYAKVFDEFPKPLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVFDEFPKPLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCCGHPEAKRMPCAEDYLSVNLQNLVLEKTPVSDRVTKCTES 480
Db 421 PTLVEVSRNLGKVGSKCCGHPEAKRMPCAEDYLSVNLQNLVLEKTPVSDRVTKCTES 480
Qy 481 LVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKEROIKQOTALVELVVKHPKAT 540
Db 481 LVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKEROIKQOTALVELVVKHPKAT 540
Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 12
US-10-413-831-1
; Sequence 1, Application US/10413831
; Publication No. US20030194813A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/413,831
; CURRENT FILING DATE: 2003-04-15
```

```
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-413-831-1

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGSENFKALVLIAPAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60
Db 1 DAHKSEVAHRFKDLGSENFKALVLIAPAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVRPEV 120
Qy 121 DVMTAFHNDNEETFLKKLYEIAHRHPFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db 121 DVMTAFHNDNEETFLKKLYEIAHRHPFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Qy 181 KLDELDRSGKASSAKQRLKCSLQKFGGERAFKAWAVARLSORFPKAEFAEVSKLVDLT 240
Db 181 KLDELDRSGKASSAKQRLKCSLQKFGGERAFKAWAVARLSORFPKAEFAEVSKLVDLT 240
Qy 241 VHTCCGHDLLCADDRAADLAKYICENQDS:SSKLKECCERPLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGHDLLCADDRAADLAKYICENQDS:SSKLKECCERPLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGFLYFYARRHPDYSVLLRLAKTYETLEK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGFLYFYARRHPDYSVLLRLAKTYETLEK 360
Qy 361 CAAADPHECYAKVFDEFPKPLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVFDEFPKPLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCCGHPEAKRMPCAEDYLSVNLQNLVLEKTPVSDRVTKCTES 480
Db 421 PTLVEVSRNLGKVGSKCCGHPEAKRMPCAEDYLSVNLQNLVLEKTPVSDRVTKCTES 480
Qy 481 LVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKEROIKQOTALVELVVKHPKAT 540
Db 481 LVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKEROIKQOTALVELVVKHPKAT 540
Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 13
US-10-413-831-2
; Sequence 2, Application US/10413831
; Publication No. US20030194813A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
```

; FILE REFERENCE: ISCO07
 ; CURRENT APPLICATION NUMBER: US/10/413,831
 ; CURRENT FILING DATE: 2003-04-15
 ; PRIOR APPLICATION NUMBER: US/09/806,247
 ; PRIOR FILING DATE: 2001-07-16
 ; PRIOR APPLICATION NUMBER: 60/115,392
 ; PRIOR FILING DATE: 1999-01-11
 ; PRIOR APPLICATION NUMBER: 60/102,738
 ; PRIOR FILING DATE: 1998-10-02
 ; PRIOR APPLICATION NUMBER: 09/165,926
 ; PRIOR FILING DATE: 1998-10-02
 ; PRIOR APPLICATION NUMBER: 09/165,581
 ; PRIOR FILING DATE: 1998-10-02
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: MOD RES
 ; LOCATION: (1)-(585)
 ; OTHER INFORMATION: ACETYLATION
 ; US-10-413-831-2

Query Match 100.0%; Score 3103; DB 12; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 DAHKSEVAHRFKDGLGEENFKALVLIATAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHKSEVAHRFKDGLGEENFKALVLIATAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60

 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVPRPV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVPRPV 120

 QY 121 DVNCTAFHNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 DB 121 DVNCTAFHNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180

 QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFAEFAEVSKLVTDLTK 240

 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

 QY 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLVEYARRHPDYSVWLLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLVEYARRHPDYSVWLLRLAKTYETTLK 360

 QY 361 CAADDPHECYAKVDFDFKPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKVPQVST 420
 DB 361 CAADDPHECYAKVDFDFKPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKVPQVST 420

 QY 421 PTLVEYSRNLGKYGSKCKHPKAPCAEDYLSVVNLQCLVHKEKTPVSDRVTKCCTES 480
 DB 421 PTLVEYSRNLGKYGSKCKHPKAPCAEDYLSVVNLQCLVHKEKTPVSDRVTKCCTES 480

 QY 481 LVNRRPCFSALEVDYTPKFEFNAFTTTHADICTLSEKERQIKKOTALVELVKKPKAT 540
 DB 481 LVNRRPCFSALEVDYTPKFEFNAFTTTHADICTLSEKERQIKKOTALVELVKKPKAT 540

 QY 541 KEQLKAVMDFFAAVFEKCKCKADDDKTCFAEEGKGLVAASQAALGL 585
 DB 541 KEQLKAVMDFFAAVFEKCKCKADDDKTCFAEEGKGLVAASQAALGL 585

RESULT 14
 US-10-153-064-5
 ; Sequence 5, Application US/10153064

; Publication No. US20020142814A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PF556
 ; CURRENT APPLICATION NUMBER: US/10/153,064
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-153-064-5

Query Match 100.0%; Score 3103; DB 14; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 DAHKSEVAHRFKDGLGEENFKALVLIATAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHKSEVAHRFKDGLGEENFKALVLIATAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60

 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVPRPV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVPRPV 120

 QY 121 DVNCTAFHNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 DB 121 DVNCTAFHNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180

 QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFAEFAEVSKLVTDLTK 240

 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

 QY 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLVEYARRHPDYSVWLLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLVEYARRHPDYSVWLLRLAKTYETTLK 360

 QY 361 CAADDPHECYAKVDFDFKPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKVPQVST 420
 DB 361 CAADDPHECYAKVDFDFKPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKVPQVST 420

 QY 421 PTLVEYSRNLGKYGSKCKHPKAPCAEDYLSVVNLQCLVHKEKTPVSDRVTKCCTES 480
 DB 421 PTLVEYSRNLGKYGSKCKHPKAPCAEDYLSVVNLQCLVHKEKTPVSDRVTKCCTES 480

 QY 481 LVNRRPCFSALEVDYTPKFEFNAFTTTHADICTLSEKERQIKKOTALVELVKKPKAT 540
 DB 481 LVNRRPCFSALEVDYTPKFEFNAFTTTHADICTLSEKERQIKKOTALVELVKKPKAT 540

 QY 541 KEQLKAVMDFFAAVFEKCKCKADDDKTCFAEEGKGLVAASQAALGL 585
 DB 541 KEQLKAVMDFFAAVFEKCKCKADDDKTCFAEEGKGLVAASQAALGL 585

RESULT 15
 US-09-984-010-7
 ; Sequence 7, Application US/09984010
 ; Publication No. US20030104578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballance, David James
 ; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
 ; AND SERUM ALBUMIN
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP

STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
MEDIUM TYPE: F-Opdy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent: Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,010
FILING DATE: 21-May-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/091,873
FILING DATE: 25-JUN-1998
APPLICATION NUMBER: PCT/GB96/03164
FILING DATE: 19-DEC-1996
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-984-010-7

Search completed: October 27, 2003, 15:17:09
Job time : 75 secs

Query Match 100.0%; Score 3103; DB 11; Length 604;
Best Local Similarity 100.0%; Pred. No. 2.5e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	DAHSEVNAHREKDLGEENFKALVLIATFAQYLCQPPFEDHVKLVNEVTEPAKTCVADESAE	60
Db	20	DAHSEVNAHREKDLGEENFKALVLIATFAQYLCQPPFEDHVKLVNEVTEPAKTCVADESAE	79
Qy	61	NCDSLSLHFGCKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDNNPLRLVRPEV	120
Db	80	NCDSLSLHFGCKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDNNPLRLVRPEV	139
Qy	121	DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP	180
Db	140	DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP	199
Qy	181	KLDELDEGKASSAKQRLKCAASLOKFGSERAKAWAVARLSORFPKAEFAEVSKLVTDLTK	240
Db	200	KLDELDEGKASSAKQRLKCAASLOKFGSERAKAWAVARLSORFPKAEFAEVSKLVTDLTK	259
Qy	241	VHTECHGDLLECDRADLAKYICENQDSISSKKECCXKPLLEKSHCIAEVENDEMPA	300
Db	260	VHTECHGDLLECDRADLAKYICENQDSISSKKECCXKPLLEKSHCIAEVENDEMPA	319
Qy	301	DLPSLAADFVESKDVCKNYAEAKVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLTKC	360
Db	320	DLPSLAADFVESKDVCKNYAEAKVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLTKC	379
Qy	361	CAAADPHCEYAKVDFEPFLVPEEPONLIKQNCCLFEQLGEYKFNQALLVRYTKVPQVST	420
Db	380	CAAADPHCEYAKVDFEPFLVPEEPONLIKQNCCLFEQLGEYKFNQALLVRYTKVPQVST	439
Qy	421	PTLVEVSNLKGKSKCKHPKEMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES	480
Db	440	PTLVEVSNLKGKSKCKHPKEMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES	499
Qy	481	LVNRRPCFSALEVDETYVPKEFNATFTFHADICTLSEKERQIKKQTALVELVXHKPKAT	540
Db	500	LVNRRPCFSALEVDETYVPKEFNATFTFHADICTLSEKERQIKKQTALVELVXHKPKAT	559
Qy	541	KEQLKAVMDDFAAFEVKCKCADDKETCFABEGKKLVAAASQAALGL	595
Db	560	KEQLKAVMDDFAAFEVKCKCADDKETCFABEGKKLVAAASQAALGL	604

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:02:02 ; Search time 43 Seconds
(without alignments)
1308.341 Million cell updates/sec

Title: US-09-832-501-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHRFKDLGEENFK.....TCFAEEGKKLVAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 76: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1 ABHUS	serum albumin prec
2	2942	94.8	602	2 A47391	serum albumin prec
3	2620	84.4	608	2 S57632	serum albumin prec
4	2475.5	79.8	607	1 ABHOS	serum albumin prec
5	2446.5	78.8	607	1 ABHOS	serum albumin prec
6	2432.5	78.4	607	1 ABHOS	serum albumin prec
7	2426	78.2	608	1 ABR7S	serum albumin prec
8	2411.5	77.7	605	1 ABPOS	serum albumin prec
9	2387	76.9	609	2 JCS838	albumin - Mongolia
10	1864	60.0	453	2 A05139	serum albumin - mo
11	1557.5	50.2	615	1 ABCHS	serum albumin prec
12	1253.5	40.4	609	2 JCA258	alpha-fetoprotein
13	1249.5	40.3	609	1 PPHU	alpha-fetoprotein
14	1242.5	40.0	609	1 PPGO	alpha-fetoprotein
15	1205	38.8	607	1 ABXL72	74K albumin prec
16	1181.5	38.1	285	2 I46986	albumin - dog (fra
17	1175.5	37.9	608	1 ABX168	68K serum albumin
18	1084	34.9	605	1 FPMS	alpha-fetoprotein
19	1067	34.4	611	1 FPRT	alpha-fetoprotein
20	1055	34.0	599	1 A54906	afamin precursor -
21	928.5	29.9	614	2 S59517	serum albumin prec
22	928	29.9	608	2 A53195	afamin precursor -
23	747.5	24.1	608	1 ABONS1	serum albumin 1 pr
24	742.5	23.9	608	1 ABONS2	serum albumin 2 pr
25	699	22.5	382	2 A37253	serum albumin - bu
26	440.5	14.2	1423	1 S27941	serum albumin - se
27	386	12.4	474	1 VYHDD	vitamin D-binding
28	385	12.4	476	1 VYRTD	vitamin D-binding
29	372	12.0	472	1 A35327	vitamin D-binding

30	184	5.9	1819	2	A71928	cag island protein
31	184	5.9	1927	2	G64585	cag pathogenicity
32	141.5	4.6	1560	2	T30282	calcium-binding pr
33	137.5	4.3	1348	2	AG2558	hypothetical prote
34	134.5	4.3	1004	2	JC2221	major surface glyco
35	134	4.3	1083	2	JC2300	cell surface glyco
36	134	4.3	1780	2	T17272	hypothetical prote
37	133.5	4.3	1390	2	S51364	sperm tail-specifi
38	132.5	4.3	1070	2	T06733	kinesin homolog P2
39	131	4.2	1076	2	JC2217	major surface glyco
40	130	4.2	1175	2	D35815	myosin heavy chain
41	129	4.2	1017	2	JC4035	cell-cycle-depende
42	128.5	4.1	1051	2	JC4091	glycoprotein A - p
43	128	4.1	1005	2	A64465	hypothetical prote
44	128	4.1	3225	2	I52300	granulin - human
45	126.5	4.1	1927	2	A59236	embryonic muscle m

ALIGNMENTS

RESULT 1

ABHUS

serum albumin precursor [validated] - human

N:Alternate names: preproalbumin

N:Contains: Kinetensin

C:Species: Homo sapiens (man)

C>Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000

C:Accession: A93743; A93936; I39427; I59286; I39427; G01747; S5314; A91420; S06422; ;
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Hock, C.M.; Najarian, R.C.; SeeI
Nucleic Acids Res. 9, 6103-6114, 1981

A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia co

A:Reference number: A93743; MUID:82081882; PMID:6171778

A:Accession: A93743

A:Molecule type: mRNA

A:Residues: 1-419; 'K', 421-609 <LAW>

A:Cross-references: EMBL:V00495; GB:J00078; GB:LC0132; GB:L00133; NID:G28591; PIDN:CAU
R:Dugaiczky, A.; Law, S.W.; Dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A:Reference number: A93936; MUID:82105994; PMID:6275391

A:Accession: A93936

A:Molecule type: mRNA

A:Residues: 1-120; 'G', 122-609 <DUG>

A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590

R:Jrncg, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

C. Biol. Chem. 261, 3244-3251, 1986

A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions at

A:Reference number: I39427; MUID:86140099; PMID:2419329

A:Accession: I39427

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-26 <URA>

A:Cross-references: GB:M13075; NID:G178330; PIDN:AAA51688.1; PID:G553173

R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian fami.

A:Reference number: I59286; MUID:94181575; PMID:8134387

A:Accession: I59286

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 282-290; 'KSRFDLQ' <WAT>

A:Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033

A>Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia

R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnar

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxy

A:Reference number: I59313; MUID:94294404; PMID:8022807

A:Accession: I59313

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 589-590; 'ALPRVRNLLQVKKLP' <NAD>

A:Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; PID:G547232

A>Note: this frame-shift variant is designated albumin Bazzano; four additional variants

R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A:Reference number: G08292
A:Accession: G01747
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-120 'G' 122-455 <MEN>
A:Cross-references: EMBL:U02961; MID:9763428; PIDN:AA64922.1; PID:9763431
Biochem. J. 308, 321-325, 1995
A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast *Kex2*
A:Reference number: S55314; MID:95275251; PMID:7755581
A:Accession: S55314
A:Molecule type: protein
A:Residues: 19-27 <LED>
R;Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A:Title: Complete amino acid sequence of human serum albumin.
A:Reference number: A91420; MID:76187907; PMID:1225573
A:Accession: A91420
A:Molecule type: protein
A:Residues: 25-117, 60', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-
R;Roehr, U.; Spittler, G.; Tripiet, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A:Title: Isolation and structure elucidation of middle-molecular weight peptides from ud
A:Reference number: S06422
A:Note: this paper is in German, with an English abstract
A:Accession: S06422
A:Molecule type: protein
A:Residues: 25-48 <ROE>
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A:Title: Mass spectrometric identification of modifications to human serum albumin treat
A:Reference number: S36882; MID:93384321; PMID:8373198
A:Accession: S36882
A:Molecule type: protein
A:Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <FIN>
R;Kausler, E.; Spittler, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A:Title: Bruchstücke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
A:Reference number: S17599; MID:92126241; PMID:1772598
A:Accession: S17599
A:Molecule type: protein
A:Residues: 25-54, 354-357, 431-447 <KAU>
A:Note: 49-Leu was also found
R;Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A:Reference number: A45800; MID:89341406; PMID:2474609
A:Accession: A45800
A:Molecule type: protein
A:Residues: 166-173, 'L' <CAR>
R;Mozard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, C.E.; Wa
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A:Title: The amino acid sequence of kinetinsin, a novel peptide isolated from pepsin-tre
A:Reference number: A03239; MID:86242180; PMID:3087352
A:Accession: A03239
A:Molecule type: protein
A:Residues: 166-173, 'L' <MOG>
R;Galliano, M.; Minichioti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madisen, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A:Title: Mutations in genetic variants of human serum albumin found in Italy.
A:Reference number: A38255; MID:91062352; PMID:2247440
A:Accession: A38255
A:Molecule type: protein
A:Residues: 76-111 <GAL1>
A:Accession: B38255
A:Molecule type: protein
A:Residues: 82-105, 'K', 107-110 <GAL2>
A:Note: this variant is designated albumin Vibo Valentia
A:Accession: A38255
A:Molecule type: protein

A:Residues: 76-83, 'K', 85-106 <GAL3>
A:Note: this variant is designated albumin Torino
R;Minichioti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A:Title: The structural characterization and bilirubin-binding properties of albumin
A:Reference number: S33298; MID:93292504; PMID:8513793
A:Accession: S33298
A:Molecule type: protein
A:Residues: 255-263, 'E', 265-281 <MIN1>
A:Note: this variant is designated albumin Herborn
R;Minichioti, L.; Galliano, M.; Scoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Por
Biochem. Biophys. Acta 1119, 232-238, 1992
A:Title: Two albumins with identical electrophoretic mobility are produced by dif
A:Reference number: S21078; MID:92190239; PMID:1347703
A:Accession: S21078
A:Molecule type: protein
A:Residues: 354-356, 'K', 358-378 <MIN2>
A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is report
R;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A:Title: Atomic structure and chemistry of human serum albumin.
A:Reference number: A46756; MID:92334427; PMID:1630489
A:Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.O.
In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-
A:Reference number: A94442
A:Contents: annotation; three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A:Title: Disulfide bonds in human serum albumin.
A:Reference number: A90930
A:Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity bindi
A:Reference number: A90299; MID:78186630; PMID:656055
A:Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
In Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-
A:Title: Serum albumin: conformation and active sites.
A:Reference number: A94408
A:Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein ge
A:Reference number: A90028; MID:83279982; PMID:6192711
A:Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic ac
A:Reference number: A46755; MID:76257808; PMID:955075
A:Contents: annotation
A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic a
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phos
A:Reference number: A56294; MID:92183881; PMID:1544460
A:Contents: annotation
A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; i
atase activity
A:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesiz
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (we
C:Comment: A large number of variants of human serum albumin have been described.
C:Genetics:
A:Gene: GDB:ALB
A:Cross-references: GDB:118990; OMIM:103600
A:Map position: 4q11-4q13
A:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrri
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-609/Product: serum albumin #status experimental <MPT>
F;29-202/Domain: serum albumin repeat homology <SA1>

F:166-174/Product: kinetensin #status experimental <KIP>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA2>
F:27/Binding site: copper (His) #status predicted
F:77-86,99,115,114-125,148-193,192-201,224-270,269-277,289-303,302-33,340-385,384-393,4
F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 9 2e-198;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAOYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDGLGENFKALVLIAPAOYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 144
QY 121 DVNCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAAFTECCQAADKAACLLP 180
Db 145 DVNCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAAFTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORFPAEFAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORFPAEFAEVSKLVTDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENODSISSKLKECCCKPLLEKSHGICAEVNDMPA 300
Db 265 VHTCCCHGDLLECCADRADLAKYICENODSISSKLKECCCKPLLEKSHGICAEVNDMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHDPYDVSVLLRLAKYETTTLEKC 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHDPYDVSVLLRLAKYETTTLEKC 384
QY 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFDFKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 540
Db 505 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 564
QY 541 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKKLVAASQAALG 595
Db 565 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKKLVAASQAALG 609

RESULT 2
A47391
serum albumin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #ext_change 20-Aug-1999
C:Accession: A47391
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwyer, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding site
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-600 <WAT>
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A:Experimental source: liver
A:Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:128281)
C:Superfamily: serum albumin; serum albumin repeat homology
F:21-194/Domain: serum albumin repeat homology <SA1>
F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>
Query Match 94.8%; Score 2942; DB 2; Length 600;
Best Local Similarity 93.5%; Pred. No. 4e-187;
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAOYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 17 DTHKSEVAHRFKDGLGEEHFKGLVLPVAFQYLOQCPFEHVKLVNEVTEFAKTCVADESAAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
Db 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 136
QY 121 DVNCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAAFTECCQAADKAACLLP 180
Db 137 DVNCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAAFTECCQAADKAACLLP 196
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORFPAEFAEVSKLVTDLTK 240
Db 197 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORFPAEFAEVSKLVTDLTK 256
QY 241 VHTCCCHGDLLECCADRADLAKYICENODSISSKLKECCCKPLLEKSHGICAEVNDMPA 300
Db 257 VHTCCCHGDLLECCADRADLAKYICENODSISSKLKECCCKPLLEKSHGICAEVNDMPA 316
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHDPYDVSVLLRLAKYETTTLEKC 360
Db 317 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHDPYDVSVLLRLAKYETTTLEKC 376
QY 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
Db 377 CAAADPHECYAKVDFDFKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 436
QY 421 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 437 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 496
QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 540
Db 497 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 556
QY 541 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKKLVAASQAAL 583
Db 557 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKKLVAASQAAL 599

RESULT 3
S57632
serum albumin precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #ext_change 20-Aug-1999
C:Accession: J04660; S57632
R:Hilger, C.; Grigioni, F.; Hentges, F.
Gene 169, 235-236, 1996
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A:Reference number: J04660; MUID:96194824; PMID:8647469
A:Accession: J04660
A:Molecule type: mRNA
A:Residues: 1-608 <H12>
A:Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485
A:Experimental source: liver
A:Comment: This protein is the major protein component in plasma. It functions as a m
ein has 35 conserved cysteine residues.
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: liver; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PP>
F:25-608/Product: serum albumin #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 84.4%; Score 2620; DB 2; Length 608;
Best Local Similarity 82.0%; Pred. No. 8.2e-166;
Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKDGBENFKALVLIAPAYVLCQCPEDHVKLVNEVTEFAKTCVADESSE 60
DB 25 EAHQSEIAHRRFNDLGEHFFGLVLVAFSQYLQCCPFEDHVKLVNEVTEFAKGCVAQDSAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPCLPRLVRPEV 120
DB 85 NCKSLHLLGCKLCTVASURDKYGENADCCCKEPERNECFLOHKDDNPGFQLVTPEA 144
QY 121 DVMTAFPHDNEETFLKYLKYLIEARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLIP 180
DB 145 DAMCTAFHENEQRFGLKYLVEIARRHPYFYAPPELLFYAEYKGVFTCECAADKAACLIP 204
QY 181 KLDELREGKASSAKQRLKASLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 205 KVDALREKVLASSAKERLKASLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 264
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISGKLKECCCKPPLLEKSHCIAEVENDEMPA 300
DB 265 IHKECCGDLLECCADDRADLAKYICENQDSISGKLKECCCKPPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEAARRHPDYVSVLLRLAKTYETTTLEKC 360
DB 325 DLPLAVDFVEDKVCNKYQAEAKDVFLGMFLYEAARRHPDYVSVLLRLAKTYETTTLEKC 394
QY 361 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFEOIDGEYKFNALLVRYTKKYPQVST 420
DB 385 CATDDPACVAVHDFEFKPLVEEPONLIKONCELFEOIDGEYKFNALLVRYTKKYPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKPKEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVEVSRNLGKVGSKCKPKEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDVTVPKFNAETFTFHADICTLSEKERQIKQOTALVELVKKPKPAT 540
DB 505 LVNRRPCFSALEVDVTVPKFNAETFTFHADICTLSEKERQIKQOTALVELVKKPKPAT 564
QY 541 KEQLKAVMDDFAAFEVCKCKADKXETCFABEGKKLVAASQAAL 583
DB 565 BEQLKTVMGDFGFSVDKCAAEKDEACFAEBEGKKLVAASQAAL 607

RESULT 4
ABHOS
serum albumin precursor - horse
C:Species: Equus caballus (domestic horse)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S34053
R:Ho, J. X.; Holowachuk, E. W.; Norton, E. J.; Twigg, P. D.; Carter, D. C.
Eur. J. Biochem. 215, 205-212, 1993
A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A:Reference number: S34053; MUID:93345495; PMID:8344282
A:Accession: S34053
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: GB:X74045; NID:G939671; PIDN:CAA52194.1; PID:G939672
A:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:11-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-Product: serum albumin #status predicted <VAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (H3s) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:77-86/Binding site: bilirubin (lys) #status predicted

Query Match 79.8%; Score 2475.5; DB 1; Length 607;
Best Local Similarity 76.3%; Pred. No. 3e-156;
Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;

QY 1 DAHSEVAHRRFKDGBENFKALVLIAPAYVLCQCPEDHVKLVNEVTEFAKTCVADESSE 60
DB 25 DTHKSEIAHRRFNDLGEHFFGLVLVAFSQYLQCCPFEDHVKLVNEVTEFAKCAADESAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPCLPRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPCLPRLVRPEV 143
QY 121 DVMTAFPHDNEETFLKYLKYLIEARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLIP 180
DB 144 DAQCAAFQEDPDKFLGKYLVEYARRHPYFYAPPELLFFAEYKADFTTECCPADDLKACLP 203
QY 181 KLDELREGKASSAKQRLKASLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 204 KLDELREGKASSAKQRLKASLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 263
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISGKLKECCCKPPLLEKSHCIAEVENDEMPA 300
DB 264 VHTCECHGDLLECCADDRADLAKYICENQDSISGKLKECCCKPPLLEKSHCIAEVENDEMPA 323
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEAARRHPDYVSVLLRLAKTYETTTLEKC 360
DB 324 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEAARRHPDYVSVLLRLAKTYETTTLEKC 383
QY 361 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFEOIDGEYKFNALLVRYTKKYPQVST 420
DB 384 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFEOIDGEYKFNALLVRYTKKYPQVST 443
QY 421 PTLVEVSRNLGKVGSKCKPKEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTES 480
DB 444 PTLVEVSRNLGKVGSKCKPKEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTES 503
QY 481 LVNRRPCFSALEVDVTVPKFNAETFTFHADICTLSEKERQIKQOTALVELVKKPKPAT 540
DB 504 LVNRRPCFSALEVDVTVPKFNAETFTFHADICTLSEKERQIKQOTALVELVKKPKPAT 563
QY 541 KEQLKAVMDDFAAFEVCKCKADKXETCFABEGKKLVAASQAAL 583
DB 564 KEQLKAVMDDFAAFEVCKCKADKXETCFABEGKKLVAASQAAL 606

RESULT 5
ABHOS
serum albumin precursor [validated] - bovine
N:Alternate names: 67K protein; preproalbumin
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
C:Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458;
R:Holowachuk, E. W.; Stoltzenberg, J. K.; Reed, R. G.; Peters Jr., T.
Submitted to the EMBL Data Library, August 1991
A:Description: Bovine serum albumin: cDNA sequence and expression.
A:Reference number: A38885
A:Accession: A38885
A:Molecule type: mRNA
A:Residues: 1-607 <HOL>
A:Cross-references: EMBL:M73215
R:Hayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
A:Title: Rapid confirmation and revision of the primary structure of bovine serum alt
A:Reference number: A36401; MUID:91083649; PMID:2260975
A:Accession: A36401
A:Molecule type: protein
A:Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>
R:MacGillivray, R. T. A.; Chung, D. W.; Davie, E. W.
Eur. J. Biochem. 98, 477-485, 1979
A:Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A:Reference number: A91258; MUID:80024278; PMID:488109
A:Accession: A91258
A:Molecule type: protein

Query Match 78.8%; Score 2446.5; DB 1; Length 607;

Best Local Similarity 75.6%; Pred. No. 2.5e-154;
Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;

Cy	1	DAHKSEVAHFRFDGLGEENFKALVLIAPAYQYLQOCFFEDHVKLWNEVTTEFAKTCVADESAE	60
Dd	25	DTHKSEIAHFRFDLGGEPKGVLVIAPSGYLOOCFFDHRVKLVNELTFEAKTCVADESHA	84
Cy	61	NCKSLHTLFGDKLTCTVATLRITYGEMADCCAKQEPRNECFLOHQDNPNLPRLVRREV	120
Dd	85	GGEKSLHTLFGDELCKVASLRETYGDMADECCEQBEPNERECFLSHKDSDSPILPKL-KFDP	143
Cy	121	DVMCTAFHDNVESTELKKLYLETARSHHPFYPAPELLFAKRYKAAATECCQAADKAACLIP	180
Dd	144	NLTCDLFKADEKKFWGKYLETARSHHPFYPAPELLYYANKYNGVFQDCQAEDEKGCALLP	203
Cy	181	KDELDRDEKASSAQRUKCASLQRFGERAFKAWAVARLSORFPKAEPAEYSKLVTDLTG	240
Dd	204	KIETMREKVLAASARQURCARCSIQFGERALKAWSVARLSOKFPKAEVEVTKLVTDLTK	263
Cy	241	VHTECCHGLLFCADRDLAKYICENODSSISKUKECCEKPLLEKSHCIAEVENDMPA	300
Dd	264	VHKECCHGLLFCADRDLAKYICDNQOTSSKUECCDKPLEEKSHCIAEVEKDAIPE	323
Cy	301	DLPSLAADVEESKVCYNZAEAKDFVLGNFEYARRHPDYSVWLILLKLAKTYETTLEKC	360
Dd	324	NUPPLTADFADKDVCKNYQEAADAFSLSFLEYSRHPPEYAVSVILLRAKEYEATLEEC	383
Cy	361	CAAADPHCYAKVFDEFKPLVBEPPNLIKQNCELFEQLGEYKFQNALIVRTTKVPQVST	420
Dd	384	CANDDPHCYSTVFQKHLVDPEPNLIKONDQDEKLGEGYGFONALIVRTRKVPQVST	443
Cy	421	PTLVEVSRLGVKGSCKKHPEAKRMPCAEEDYLSVLNQLCVLHEKTPVSORVTKCCCTES	480
Dd	444	PTLVEVSRSJGVKGTCTCKPSESNPCTEDYLSLILNRCLCVLHEKTPUSEKVTACCTES	503
Cy	481	LVNRRPCFSALEVDITYVPKEFNATFTFHADICTLSEKERQKKQTALVELVGHKPKAT	540
Dd	504	LVNRRPCFSALTDETYPVPKAFDEKLFTFHAGICTLPOTEKQKKQTALVELLVGHKPKAT	563
Cy	541	KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAL	583
Dd	564	EBOCLKTMENFVAFVDKCCAADDKEACTFAVEGKLVVSTQIAL	606

RESULT 6
ABSHS
serum albumin precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S06936
R:Brown, W.X.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A:Reference number: S06936; MUID:90098888; PMID:2602160
A:Accession: S06936
A:Molecule type: mRNA
A:Residues: 1-607 <BRO>
A:Cross-references: EMBL:X17055; NID:g1386; PIDD:CAA34903.1; PIDD:g1387
C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the mem-
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276
F:263/Binding site: bilirubin (lys) #status predicted


```
Best Local Similarity 75.0%; Pred. No. 2,1e-153;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHKSEVAHRFDKLGSENFKALVLIAPAYQYLOQCDFEDHVKLVNEVTEFAKTCVADESAAE 60
D 25 DTHKSEIAHRFDKLGSENFKALVLIAPAYQYLOQCDFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDNDPNLPLVREPEV 120
D 85 GCDKSLHTLFGDELCKVATLRETYGDMDCCKQEPERNECFLOHKDNDPNLPLVREPEV 143
QY 121 DVNCTAFHNDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCQADKAACLLP 180
D 144 DTLCAEFKADKFKGKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCQADKAACLLP 203
QY 181 KLDELRDEGKASSAKRLKCSLQKGFGERAFKAMAVARLSQRFKAEFYAEVSKLVTLTK 240
D 204 KIDAKREKVLASSAKQRURCSLQKGFGERALKAMSVARLSQRFKADFTDVKIATVTLTK 263
QY 241 VHTCCHGDLLECCADDRADLAKYICENQDS:SSKLKECCCKPLLEKSHCAEVENDEMPA 300
D 264 VHKCCHGDLLECCADDRADLAKYICDHQDALSSKLKECCCKPLLEKSHCAEVENDEMPA 323
QY 301 DLPSLAADPVEKSKYAEKDYFLGMFLYEAARRHPYFYAPPELLFFAKRYKAAAFTECCQADKAACLLP 360
D 324 NLPLPTADPAEDKVEKCKYAEKDYFLGMFLYEAARRHPYFYAPPELLFFAKRYKAAAFTECCQADKAACLLP 383
QY 361 CAADDPHECYAKVDFEFLVPEEPQNLKONCELFGEQYKFNQALVRYTKVQVQVST 420
D 384 CAKEDPHACYATVDFKLKHLVDEPNLKKONCELFGEQYKFNQALVRYTKVQVQVST 443
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
D 444 PTLVEISRLGKVGKCCAKPESERMECTEDYLSVLNQLCVLHEKTPVSEKVTCTES 503
QY 481 LVNRRPCFSALEVDETYVPKPENAETFFHADICTLSEKEROIKQTALVELVGHKPKAT 540
D 504 LVNRRPCFSDLTLDCTVYVPKPEDEKFFTHADICTLPDTEKQIKQTALVELVGHKPKAT 563
QY 541 KEQLKAVMDDFAAVEKCKCKADKCTCAEAGCKLVAAASQAL 583
D 564 DEQLKXTVMENFVAFVDDKCAADKCKGCFVLEGFKLVASTQAL 606

RESULT 7
ABRIS
serum albumin precursor - rat
N:Alternate names: preproalbumin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 22-Jun-1999
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:81223722; PMID:7017712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <SAR>
A:Cross-references: GB:J00698; NID:G55627; PIDN:CAA24532.1; PID:G55628
R:Stauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A:Reference number: A92211; MUID:77249657; PMID:893447
A:Note: Cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <STR>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein
```

```
A:Residues: 25-222 <ISI>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino
A:Reference number: A91940; MUID:76260153; PMID:956149
A:Accession: A91940
A:Molecule type: protein
A:Residues: 223-288:572-608 <IS2>
A:Note: 262-Leu was also found
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
A:Reference number: A90758; MUID:79001617; PMID:80265
A:Contents: annotation; copper binding
R:Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1890-1898, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid prote
A:Reference number: A45800; MUID:89341406; PMID:2474609
A:Accession: C45800
A:Status: preliminary
A:Molecule type: protein
A:Residues: 166-173 <CAR>
R:Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an impro
A:Reference number: I57621; MUID:87286876; PMID:3475566
A:Accession: I57621
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: GB:M16825; NID:G202828; PIDN:AAA40712.1; PID:G554412
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status experimental <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-608/Product: serum albumin #status experimental
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (HIS) #status experimental
F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-385
Query Match 78.2%; Score 2426; DB 1; Length 608;
Best Local Similarity 73.4%; Pred. No. 5.6e-153;
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFDKLGSENFKALVLIAPAYQYLOQCDFEDHVKLVNEVTEFAKTCVADESAAE 60
D 25 EAKSEIAHRFDKLGSENFKALVLIAPAYQYLOQCDFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDNDPNLPLVREPEV 120
D 85 NCDKSLHTLFGDKLCAIPKLRDNYGELADCCAKQEPERNECFLOHKDNDPNLPLVREPEV 144
QY 121 DVNCTAFHNDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCQADKAACLLP 180
D 145 EAMCTSFQENPTSLFGLHYLHEVARRHPYFYAPPELLFFAKRYKAAAFTECCQADKAACLLP 204
QY 181 KLDELRDEGKASSAKRLKCSLQKGFGERAFKAMAVARLSQRFKAEFYAEVSKLVTLTK 240
D 205 KLDVAVEKALVAARVQRKCSMQRFGERAFKAMAVARLSQRFKAEFYAEVSKLVTLTK 264
QY 241 VHTCCHGDLLECCADDRADLAKYICENQDS:SSKLKECCCKPLLEKSHCAEVENDEMPA 300
D 265 INKCCCHGDLLECCADDRADLAKYICENQDS:SSKLKECCCKPLLEKSHCAEVENDEMPA 324
QY 301 DLPSLAADPVEKSKYAEKDYFLGMFLYEAARRHPYFYAPPELLFFAKRYKAAAFTECCQADKAACLLP 360
D 325 DLPSTAADPVEKSKYAEKDYFLGMFLYEAARRHPYFYAPPELLFFAKRYKAAAFTECCQADKAACLLP 384
QY 361 CAADDPHECYAKVDFEFLVPEEPQNLKONCELFGEQYKFNQALVRYTKVQVQVST 420
D 385 CAEGDPPACTVGLAEFQPLVEEPQNLKONCELFGEQYKFNQALVRYTKVQVQVST 444
```

Db 262 VHKECCHGDLLECCADDRADLAKYICENQDITSTLKKECCDKPLLEKSHCIAEAKDELP 321

Qy 301 DLPSLAADFVESKDVCKNRYAEAKDVFLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLKCC 360

Db 322 DLNPLEHDFVEDKEVCCKNRYEAKDVFLGTFLYFYSRRHPDYSVSLLLRLLAKTYEATLEDC 381

Qy 361 CAAADPHCEYAKVDFDEPKPLVBEPPONLIKONCELFEOQLGEYKFNALLVRYTKVPQVST 420

Db 382 CAKEDEPPACVATVDFKFPQPLVDEPPNLIKONCELFEXLGEYGFONALVRYTKVPQVST 441

Qy 421 PT:VEVYSRNLGKVGSKCKKHPKAPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480

Db 442 PTUVEVARKLGLVGSRCCKPPEERLSCAEDYLSVLNRLCVLHEKTPVSKVTKCCTES 501

Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVLVLRHKPRAT 540

Db 502 LVNRRPCFSALTPTDETYKPEFVEGTFTFHADLCTLPDEKQIKKQATLVELLRHKPHAT 561

Qy 541 KEOLKAVMDFFAPFVEKCCXADDKBTCEAEBGKLV 576

Db 562 EQQLRTVLGNFAAPVQKCAAPDEACFAVEGPKFV 597

RESULT 9

JC5838

albumin - Mongolian jird

C:Species: Meriones unguiculatus (Mongolian jird)

C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000

C:Accession: JC5838

R:Yoshida, K.; Seto-Ohshima, A.; Sinohara, H.

DNA Res. 4, 351-354, 1997

A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in

A:Reference number: JC5838; MUID:98116663; PMID:9455485

A:Accession: JC5838

A:Molecule type: mRNA

A:Residues: 1-609 <YOS>

A:Cross-references: DDBJ:AB006197; NID:G2317277; PIDN:BAA21765.1; PID:G2317278

A:Experimental source: liver

C:Superfamily: serum albumin; serum albumin repeat homology

F:222-395/Domain: serum albumin repeat homology <SA2>

Query Match 76.9%; Score 2387; DB 2; Length 609;

Best Local Similarity 73.9%; Pred. No. 2.1e-150;

Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;

Qy 2 AHKSEVAHRPKDGEENFKALVLIAFQYLOCCFDEHVKLNVETSPAKTCVADESAEN 61

Db 27 AHKSEIAHRYKDLGERYFKGLVLVTFQYLOKCSYEEHVKLVRVETDFASNCADKDESAEN 86

Qy 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDZNPRLVLRPEVD 121

Db 87 CDKSLHTLFGDKLCSLNFGEKYAEAMADCCAKQSPERNECFLOHKDDNPQLPFFKRAEPD 146

Qy 122 VMCTAFHDNEETPLKYLKLYEIAARRHPFYAPPELLFFAKRYKAAFTCECCQAADKAACLPK 181

Db 147 AMCTAFQENAEAFMGHYLHEVARRHPFYGPELLVLADKYTAVLTCECAADCKGACLPK 206

Qy 182 LDELDRDGKASSAKRKLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTLTIKV 241

Db 207 LDALKALVSAVRQKLCSSMKFGERAFKANAVARMSQTFPNADFAEITKJATLTIKV 266

Qy 242 HTECCHGDLLECCADDRADLAKYICENQDSTSSKLKECKEPLEKSHCIAEAVENDEMPAD 301

Db 267 TQECCHGDLLECCADDRADLAKYKMCENQASISSKLQACDCKEMLOKSOCLAEVEHDDMPAD 326

Qy 302 LPSLAADFVESKDVCKNRYAEAKDVFLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLKCC 361

Db 327 LPALTADPVEDKDVCKNRYAEAKDVFLGTFLYFYSRRHPDYSVSLLLRLLAKTYEATLEKCC 386

Qy 362 AAADPHCEYAKVDFDEPKPLVBEPPONLIKONCELFEOQLGEYKFNALLVRYTKVPQVSTP 421

Db 387 AEADPHACYGHVDFDEPKPLVBEPPONLVKSNCELYEKUGYGFONALVRYTKAPQVSTP 446


```

Db 330 LPSLVKVIYEDKVECKSPAGHAFMAEFVYVSRHPEFISQIMRIAKGYESLLEKCC 389
QY 362 AAADSEHCYAKVDFEKLPLVEEPQNLIKQNCLEFQGLGEYKFNQALLVRYTKKVPQVSTP 421
Db 390 KTDNPAECYANAQBNLQHIKETQDVWNTGDLHDGEADFJSLIRYTKKVPQVPTD 449
QY 422 TLVEYSRLNGLKVGSKCKHPKAPCAEDYLSVVLNQCVLHKEKTPVSORTVKCTTESL 481
Db 450 LLETGKQWTTIGTKCCOLGEDRRMACSEGYLSVIHETCRKQETTPINDNVSCCSQLY 509
QY 482 VNRPCPSALEVDETYVPKFNATFTTHADICTLSEKQIKKQKOTALVELVHKPKATK 541
Db 510 ANRRPCFTAMGVDTKYVPPENPOMFSEDEKLCAPAEEREVGQMKLINLIKRPQWTE 569
QY 542 EOLKAVMDDEAFVEKCKKADKSTCFPAEKGKLVAAASQAALGL 585
Db 570 EQIKTIADGFTAMVQKCKQCKDINTCFGEEGANLIVOSRATLGI 613

RESULT 12
JC4258
alpha-fetoprotein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C>Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C:Accession: JC4258
R:Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczky, A.
Gene 162, 213-220, 1995
A:Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to
A:Reference number: JC4258; MUID:96032345; PMID:7557431
A:Accession: JC4258
A:Molecule type: DNA
A:Residues: 1-609 <NIS>
A:Cross-references: GB:U21916; NID:9841311; PIDN:AAA91641.1; PID:9841312
C:Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
C:Similar properties and structure.
C:Genetics:
A:Gene: afp
A:Map position: 3p
A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551/
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-609/Product: alpha-fetoprotein #status predicted <MA>
F:29-202/Domain: serum albumin repeat homology <SAL>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:442,251/Binding site: carbohydrate (Asn); (cova.ent) #status predicted

Query Match 40.4%; Score 1253.5; DB 2; Length 609;
Best Local Similarity 40.1%; Pred. No. 2e-75;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;
QY 3 HKSE-----VAHRFKDLGEENFKALVLTAFAQYLCQCFPEHVKLVNEVTFATCVADE 57
Db 22 HRNEYGASILDVSOCTAEINLTDLTAFFPAQVQEAITYKEVSKWVKDALTAIEKPGDGE 81
QY 58 SAENCDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQEPNECFLOHKOONP-NLRLV 116
Db 82 QSAGCLENQLAFLEBELCREKEILEKYGH-SQCCSQSEGRHNCFLAHKQTPPASIFFQ 140
QY 117 RPEVDMCTAFHDNETELKKLYEIAARRHYFYVAPPELLFFAKRYKAAFTCCQAADKAA 176
Db 141 VPEPTVSCAEVEEDRETPMNFVIEIARRHFLYAPTILLWAARYCKIIPSCCKAENAVE 200
QY 177 CLLPKLDLDEDEGKASSAKQRIKCNLSQKQFGERAFKAWAVARLSORPKAEFAVSKLV 236
Db 201 CFQTKAAATVTKELRESSLNGHACAVMNFGRTFQAITVTKLSQKFTKNFTFIQKLVL 260
QY 257 DLTQVHTTCCCHDLLECADRADLAKYICENODSTSSKLECKCEKPLLEKSHCIAEVND 296
Db 261 DVANVHEHCCRGDVLJDCQDEKIMSYICSOQDITLSNKITTECKLTTLERQCQIIHAEND 320
QY 297 EMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYVARRHPDYSVVLLRLAKTYETT 356

```

```

Db 321 ERPEG:SPNLNRFGLDRDFNQFSSGKKNIFLASFVHEYSRRPQLAVSVILRVAKGYQEL 380
QY 357 LEKCCAAADPHCYAKVDFEKLPLVEEPQNLIKQNCLEFQGLGEYKFNQALLVRYTKKVP 416
Db 381 LEKCFOTENPLECDQKGBEELQKYIQESQALAKRSCGLFKLGEYVQLQNAFLVAVTTKKAP 440
QY 417 QVSTPTLVEVSRNLKVGSKCKHPKAPCAEDYLSVVLNQCVLHKEKTPVSORTVKTC 476
Db 441 QUTSSSELMAITTKMAATAATCCOLSEDKLLACGEGAADIIIGHLCIRHETTPVNPVGQCC 500
QY 477 CTESLVNRPCPSALEVDETYVPKFNATFTTHADICTLSEKQIKKQKOTALVELVHKH 536
Db 501 CTSSYANRRPCPSLVLVDETYVPPAFSDDKPFIPKDLCOAQGVALCTMKQEFLLNLVKOK 560
QY 537 PKATKEQLKAVMDDEAFVEKCKKADKSTCFPAEKGKLVAAASQAALGL 585
Db 561 POTTERQLEAVIADFSGLLEKCKQEQEVCFPAEKGKLVAAASQAALGLV 609

```

RESULT 13

PFHU

```

alpha-fetoprotein precursor [validated] - human
N:Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
C:Accession: A26624, S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042;
R:Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczky, A.
Biochemistry 26, 1332-1343, 1987
A:Title: Structure, polymorphism, and novel repeated DNA elements revealed by a compl
A:Reference number: A26624; MUID:87185438; PMID:2436661
A:Accession: A26624
A:Molecule type: DNA
A:Residues: 1-609 <GB>
A:Cross-references: GB:M1610; NID:9773678; PIDN:AA585754.1; PID:g178236
R:McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tighman, S.; Krum
Hum. Mol. Genet. 2, 379-384, 1993
A:Title: A G->A substitution in an HNF I binding site in the human alpha-fetoprotein
A:Reference number: S37655; MUID:93278385; PMID:7684942
A:Accession: S37655
A:Molecule type: DNA
A:Residues: 1-28 <MCV>
A:Cross-references: EMBL:219532; NID:928527; PIDN:CAA79592.1; PID:g28528
A:Note: the authors translated the codon TAT for residue 26 as Thr
R:Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A:Title: Primary structures of human alpha-fetoprotein and its mRNA.
A:Reference number: A93961; MUID:83273664; PMID:6192439
A:Accession: A93961
A:Molecule type: mRNA
A:Residues: 1-609 <MOR>
A:Cross-references: GB:J00077; NID:g311348; PIDN:CAA24758.1; PID:g31351
R:Beattie, W.G.; Dugaiczky, A.
Gene 20, 415-422, 1982
A:Title: Structure and evolution of human alpha-fetoprotein deduced from partial seq
A:Reference number: A91497; MUID:83158778; PMID:6187626
A:Accession: A91497
A:Molecule type: mRNA
A:Residues: 429-556 <BEA>
A:Cross-references: GB:J00076
R:Pucci, P.; Siciliano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Teri
Biochemistry 30, 5061-5066, 1991
A:Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A:Reference number: A23699; MUID:91242409; PMID:1709810
A:Accession: A23699
A:Molecule type: protein
A:Residues: 19-45;60-97;102-107;122-184;187-249;255-489;507-609 <PUC>
R:Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A:Title: Characterization of in vitro expressed human alpha-fetoprotein as highly ref
A:Reference number: A61480; MUID:91225826; PMID:1709209
A:Accession: A61480
A:Molecule type: protein

```

Query Match	40.3%	Score 1249.5;	DB 1;	Length 609;
Best Local Similarity	39.9%;	Pred. No. 3.8e-75;		
Matches 235; Conservative 116;	Mismatches 231;	Indels 7;	Gaps 3;	

QY	3	HKSE-----VAHAFKOLGEENFKALVLIAFAQLVQCQP	FEDHVKLINVEVTEFAKTCAUAE	57
		:::::	:::::	:::::
Dd	22	HRNEYGIASILDSVQCCTAESLADLATIFPAFVQEATYKEVSKWVKDALTA-EKPTGDE	81	
		:::::	:::::	:::::
QY	58	SANCKDKSLHTLGCKLCTVATTRETYGEMADCCAKQEPERNECFLOHQDDNP-NAPRLV	116	
		:::::	:::::	:::::
Dd	82	QSSGCLENOLUPALEELCHEKEILELYKH-SDCCSQSEEGRHNCFLAHKKPTPASIPJFC	140	
		:::::	:::::	:::::
QY	117	RPEVDVMTAFHNEETFLKKYLIEYIARRHIPPIYABELFFAKRYKAATTECCQAADKAA	176	
		:::::	:::::	:::::

Db 141 VPEPVTSCEAYESDRETFMKNFYEIARRHPFLYAPITLWAAARYDKLIIPSCCKAEAV 200

Qy 177 CLLPKLDELDEGKASSAKORLKCASLOKGTGEGRAKAWAVARLSQRFPKPAEAEVSKLVT 236

Db 201 CFCTKAATVTKELRESSLLNQHACAVMKNFGTTFQAITVTKLSQKFTKVNFTFQKLV 260

Qy 237 DLTKVTECHGDLLECADRADLAKYICENQDSISKLEKCEKFLLEKSKHICIAEVEND 296

Db 261 DVAHVHEHCRCGVDLCLQDGEKIMYSICQDQTSNKITECCCKLTTLERGQCIHAEND 320

Qy 297 EMPADLPSLAADPVESKDVCKNVAEAKDVFLGFLMFLYEARHPDYSVVLRLRLAKYET 356

Db 321 EKPEGLSPNLNRFLGDRDFNQFSGEKNIELASFVHEYSRRHPQLAVSVILRVAKGYQL 380

Qy 357 LEKCCAAADPHECVAKVDFEFKPLVEBPQNLIKONCELFEOUGEYKFNQALLVRYTKKVP 416

Db 381 LEKCFOTENPLEQDKEEELQYIQESQALAKRSCLFQKLEYLQNAFLVAYTKKAP 440

Qy 417 QVSTPTLVEVSRNLGKVGSKCKHPKAEKRMPCAEVYLVVLNQLCVILHEKTPVSDRTKC 476

Db 441 QLTSSLEMAITRKMAATAATCCQLSDKLLACGEGRAADIIGHLCIRHEWTPVNPVGQC 500

Qy 477 CTSSLNRRPCFSALEVDETVYVPEFNAETFTTHADICTLSEKERQIKKGTALVELVKHK 536

Db 501 CTSSYANRRPCFSLVVDETVYVPAFSDDKFIHFKOLCQAGVALQTMKQEFNLINLVKQ 560

Qy 537 PKATKEOLKAVMDDFAAFVEKCKKADDKETCFPAEEGKKLVAAASQAALGJ 585

Db 561 PQITEBQLETVIADFFSGLEKCCQGGQEVCFPAEEGQKLISKTRTALGV 609

RESULT 15

ABXL72

74K albumin precursor - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text-change 22-Jun-1999

C:Accession: B41682; S02693; A05288

R:Moskatis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R. Mol. Endocrinol. 3, 464-473, 1989

A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.

A:Reference number: A41682; MUID:89313708; PMID:2747653

A:Accession: B41682

A:Molecule type: mRNA

A:Residues: 3-607 <MCS>

A:Cross-references: GB:M21442; NID:g2133930; PIDN:AAA49637.1; PID:g213391

R:Schorpp, M.; Doebebling, U.; Wagner, U.; Ryffel, G.U. J. Mol. Biol. 199, 83-93, 1988

A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deletion during development.

A:Reference number: S02692; MUID:88172470; PMID:2451026

A:Accession: S02693

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-48 <SCH>

A:Cross-references: EMBL:Z26826

R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J., Eur. J. Biochem. 146, 489-496, 1985

A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilization of the mRNA.

A:Reference number: A05288; MUID:85126974; PMID:3971963

A:Accession: A05288

A:Molecule type: mRNA

A:Residues: 459-502, 'L', 504-557 <WOL>

A:Cross-references: GB:M28276

A:Note: the authors translated the codon TAT for residue 63 as Thr

C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thyroglobulin, and retinol.

C:Genetics:

A:Introns: 27/1

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma

F:1-18/Domain: signal sequence #status predicted <Sig>

F:19-24/Domain: propeptide #status predicted <PRO>

```

F:25-607/Product: 74X serum albumin #status predicted <MAT>
F:32-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:330/Binding site: copper (His) #status predicted
F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,298-302,301-312,339-364,383-393/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          38.8%; Score 1205; DB 1; Length 607;
Best Local Similarity 39.3%; Pred. No. 3 3e-72;
Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;

QY 3 HKSEVAHRFKDGEENFKALVIAFAQYQQQPFEDHVKLVEVTFBFACTVADESAEKC 62
Db 30 HHKRIADVVTALRTFKGLTLAVSONLQKCSLELSKLVINDFAXSCNDKTPC-C 88
QY 63 DKSLSLTFGDKLCTVAITRETYGEMADCAKQPEPERNECFLOHKDONPNLPLRVREVDV 122
Db 89 EKPVGTLTFDKLCADPAVGWYVSEWSECCAKQDPERACCFKAHRDHEHT---SKPEPE 145
QY 123 MCTAFHDNNEETFLKKLYELIARHPVYFAPELLFFPAKRYKAAFTCCQAADKAACLLPKL 182
Db 146 TCKLJLKHPPDDLLSAIHEARNHPOLYPVAVUALTKQYHKAEHCCEDEKCESEKM 205
QY 193 DELRDEGKASSAKORLIKASLOKGERAFKAMAVARLSORFPKAEPAEVSCLKVTDITKVH 242
Db 206 KQLMKQSHSTEDKQHFHCWILDNFPSEKVLKALNARVSHRYPKABFKLAHNFTFEVTHFI 265
QY 243 TECHGDLLFECADDRADLAKYICENODSSSKLKECEKXPJLEKSHCIAEVENDEMPADL 302
Db 266 KDCCHDDMECFMTERLELTHEHTCOHKDELSSLKKECCNIPJLEERTYCVITJENDDDVPAEL 325
QY 303 PSLAADPFVESKCYKNYAKADVFLGWFLEYEARRHDPYSVULLLLAKTYETTLSEKCCA 362
Db 326 SQPITFTDPHVCYKAYENNEVFLGRYLHAYSRKHQELSEQFLLQSAKEYESLNNCKCK 385
QY 363 AADPHCYAKVDFDKPLVEEPPONLIKQNCLEFQELGKEYKFNQALLVRYTKVPQVSTPT 422
Db 386 TDNPPECYKDGAIRFWEAKERFAYLKQNCDLIHEHGEVLFENELLIRVTKKMPQVSDET 445
QY 423 LVEVSRNLGKVGSKCKHBEAKEMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCCTESLV 482
Db 446 LIGIAHQMADIGEHCAVPEQNOMPCAEGDLTILIGMCEKQKKTFINNHVAHCCDTSYS 505
QY 483 NRRPCTSALEVDVTVYPKFNAEFTTFHADICTLSEKERQKKQATALVELVKKHKPRATKE 542
Db 506 GMRSCFTALGPDYVPPVPPVTDTFHDDKICTANDKEKQHIKQKEVLKLIKVPSPLEKN 565
QY 543 QLKAVMDDPAAFFVEKCKKADDKETCTFAEBGKKLVAAQS 580
Db 566 HIDECSAEFLKWKQCTADCHQPCDFTEKPVJ--EHCO 603

```

Search completed: October 27, 2003, 15:07:59
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 14:54:42 ; Search time 26 Seconds
(without alignments)

1058.101 Million cell updates/sec

Title: US-09-832-501-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHRFKDLEENFK.....TCFAEEGKLVAAQAALGL 565

Scoring table: ELOSUM62

Gapop 10.C , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1	ALBU_HUMAN
2	2942	94.8	600	1	ALBU_MACMC
3	2620	84.4	608	1	ALBU_FELCA
4	2562	82.6	608	1	ALBU_CANFA
5	2475.5	79.8	607	1	ALBU_HORSE
6	2450.5	79.0	607	1	ALBU_BOVIN
7	2446	78.8	608	1	ALBU_RABIT
8	2432.5	78.4	607	1	ALBU_SHEEP
9	2426	78.2	608	1	ALBU_RAT
10	2411.5	77.7	605	1	ALBU_PIG
11	2387	76.9	609	1	ALBU_MERUN
12	2378	76.6	608	1	ALBU_MOUSE
13	1557.5	50.2	615	1	ALBU_CHICK
14	1253.5	40.4	609	1	FETA_PANTR
15	1249.5	40.3	609	1	FETA_HUXAN
16	1242.5	40.0	609	1	FETA_GORGO
17	1205	38.8	607	1	ALB2_XENLA
18	1200	38.7	609	1	FETA_HORSE
19	1164.5	37.5	606	1	ALBL_XENLA
20	1084	34.9	605	1	FETA_MOUSE
21	1067	34.4	611	1	FETA_RAT
22	1055	34.0	599	1	AFAM_HUMAN
23	944	30.4	611	1	AFAM_MOUSE
24	928	29.9	608	1	AFAM_RAT
25	747.5	24.1	608	1	ALBL_SALSA
26	742.5	23.9	608	1	ALB2_SALSA
27	699	22.5	382	1	ALBU_RANCA
28	440.5	14.2	1423	1	ALBU_PETMA
29	386	12.4	474	1	VTDB_HUMAN
30	381	12.3	476	1	VTDB_RAT
31	378	12.2	476	1	VTDB_RABIT
32	372	12.0	472	1	VTDB_MOUSE
33	151.5	4.9	1605	1	RRLB_MOUSE

34	144.5	4.7	8797	1	SNBL_HUMAN	Q8rf91, homo sapien
35	138.5	4.5	1410	1	RRLB_HUMAN	Q9p2e9, homo sapien
36	133.5	4.3	1391	1	MST2_DROHY	Q8696 drosophila
37	132.5	4.3	2230	1	GOG4_HUMAN	Q13439 homo sapien
38	129.5	4.2	1972	1	MYHB_HUMAN	P35749 homo sapien
39	129	4.2	3210	1	CNFB_HUMAN	P49454 homo sapien
40	128	4.1	1005	1	RA50_METUA	Q58718 methanococc
41	126.5	4.1	1972	1	MYHB_RABIT	P35748 oryctolagus
42	126	4.1	1189	1	SMC2_CHICK	Q90888 gallus gall
43	126	4.1	3259	1	GIAN_HUMAN	Q24789 homo sapien
44	125	4.0	3038	1	TRIO_HUMAN	Q75962 homo sapien
45	124.5	4.0	1790	1	USO1_YEAST	P25386 saccharomyc

ALIGNMENTS

RESULT 1
ALBU_HUMAN STANDARD; PRT: 609 AA
AC P02768: O95574; Q13140; Q9P157; Q9P177; Q9CHS3; Q9ULZ0;
DT 21-JUL-1986 (Rel. 01, Created)
CT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_SEQUENCE FROM N.A. MEDLINE=86196112; PubMed=3009475;
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczky A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]_SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=82081882; PubMed=6171778;
RA Law R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Rajarajan R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]_SEQUENCE FROM N.A., AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczky A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]_SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]_SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
RX TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [6]_SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RX Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [7]_SEQUENCE FROM N.A.

- RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8].
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravek L., Koska V.;
RT "Complete amino acid sequence of human serum albumin";
RL FEBS Lett. 58:134-137(1975).
RN [9].
RP SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.Q.;
RL (In) Bing D.H. (eds.);
RT The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10].
RP SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Meraya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11].
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12].
RP SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
acetylsalicylic acid";
RL FEBS Lett. 66:173-175(1976).
RN [13].
RP SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [14].
RP DISULFIDE BONDS.
RA Saber M.A., Stockbauer P., Moravek L., Meloun B.;
RT "Disulfide bonds in human serum albumin";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15].
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
affinity binding of bilirubin";
RL Biochem. J. 171:453-459(1978).
RN [16].
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
domain of serum albumin";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17].
RP VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=88068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
Amerindian and Japanese populations";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18].
RP VARIANTS NAG-1, HIR-1, HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita Y.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese alloalbumins";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19].
RP VARIANTS MANAJS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20].
RP DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
human serum albumin whose precursor has an aberrant signal peptidase
C-cleavage site";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21].
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Mirchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
Italy";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22].
RP VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Mirchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
carboxyl-terminal variants of human serum albumin";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23].
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RX MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Amaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24].
RP VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
albumin: albumin Casebrook (494 Asp-->Asn)";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25].
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two alloalbumins with identical electrophoretic mobility are produced

DE Serum albumin precursor (Allergen Can # 3).

GN ALB.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID:9615;

RN (1)

RN SEQUENCE FROM N.A.

RC STRAIN=Beagle; TISSUE=Liver;

RA Hilger C.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN (2)

RN SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=20148667; PubMed=15669843;

RA Pandjaitan B., Soboda I., Brandesky-Pichler F., Rumpold H.,

RA Valenta R., Spitzauer S.;

RT "Escherichia coli expression and purification of recombinant dog

RT albumin, a cross-reactive animal allergen.";

RL J. Allergy Clin. Immunol. 105:279-285(2000).

RN (3)

RN SEQUENCE OF 25-48.

RX MEDLINE=7501422; PubMed=4414013;

RA Dixon J.W., Sarkar B.;

RT "Isolation, amino acid sequence and copper(II)-binding properties of

RT peptide (1-24) of dog serum albumin.";

RL J. Biol. Chem. 249:5872-5877(1974).

RN (4)

RN SEQUENCE OF 25-38.

RC TISSUE=Heart;

RX MEDLINE=98163340; PubMed=9504812;

RA Dunn M.J., Corbett J.M., Wheeler C.H.;

RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of

RT dog heart proteins.";

RL Electrophoresis 18:2795-2802(1997).

RN (5)

RN SEQUENCE OF 215-478 FROM N.A.

RC TISSUE=Salivary gland;

RX MEDLINE=94201492; PubMed=7512102;

RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,

RA Muehl S., Ebner C., Schiner O., Kraft D., Rumpold H.;

RT "Molecular characterization of dog albumin as a cross-reactive

RT allergen.";

RL J. Allergy Clin. Immunol. 93:614-627(1994).

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; AJ133489; CAB64867.1; -

DR EMBL; Y17737; CAAY76841.1; -

DR EMBL; S72946; AAB30434.1; -

DR HSSP; P02768; 1E7B

DR HSC-2DPAGE; P49822; DOG.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; transport prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; Albumin; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.

FT	SIGNAL	1	18	POTENTIAL.
FT PROPEP	19	24		
FT CHAIN	25	608		SERUM ALBUMIN.
FT DOMAIN	25	205		ALBUMIN 1.
FT DOMAIN	212	397		ALBUMIN 2.
FT DOMAIN	404	595		ALBUMIN 3.
FT METAL	27	27		COPPER (BY SIMILARITY).
FT DISULFID	77	86		BY SIMILARITY.
FT DISULFID	99	115		BY SIMILARITY.
FT DISULFID	114	125		BY SIMILARITY.
FT DISULFID	148	193		BY SIMILARITY.
FT DISULFID	192	201		BY SIMILARITY.
FT DISULFID	224	270		BY SIMILARITY.
FT DISULFID	269	277		BY SIMILARITY.
FT DISULFID	289	303		BY SIMILARITY.
FT DISULFID	302	313		BY SIMILARITY.
FT DISULFID	340	385		BY SIMILARITY.
FT DISULFID	384	393		BY SIMILARITY.
FT DISULFID	416	462		BY SIMILARITY.
FT DISULFID	461	472		BY SIMILARITY.
FT DISULFID	485	501		BY SIMILARITY.
FT DISULFID	500	511		BY SIMILARITY.
FT DISULFID	538	583		BY SIMILARITY.
FT DISULFID	582	591		BY SIMILARITY.
FT CONFLICT	1	26		MKWTFISLFFLFSSAYSRGLVRREA -> MDT (IN REF. 2)
FT CONFLICT	146	146		A -> R (IN REF. 2)
FT CONFLICT	206	206		I -> T (IN REF. 2)
FT CONFLICT	349	349		V -> A (IN REF. 2)
FT CONFLICT	359	359		S -> A (IN REF. 1)
FT CONFLICT	448	448		V -> VV (IN REF. 1)
FT CONFLICT	474	474		D -> E (IN REF. 1)
SQ SEQUENCE	608 AA; 58606 MW; 3CF1C8FF7DD8FC06 CRC64;			

Query Match 82.6%; Score 2562; DB 1; Length 608;

Best Local Similarity 79.8%; Pred. No. 2.4e-159;

Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;

Qy	1	DAKSEVAHRFKDLGSENFKAFLVIAFAOYLQOCDFEDHVKLVNEVTSFAKTCVADESAAE 60
Db	25	EAYKSEIAHYRNDLGEHFRGVLVAFSQYLQOCDFEDHVKLVNEVTSFAKTCVADESAAE 84
Qy	61	NDCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPINLPRLVRPEV 120
Db	85	NDCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPINLPRLVRPEV 144
Qy	121	DWCTAFHNEBTFLLKKYLYEYARRHPYFYAPPELLFFAKRYKAAFTCCQAAKACLLP 180
Db	145	DALCAAFQDNEQLFLGKLYEYARRHPYFYAPPELLFYAQYKGVFAECQAAKACLLP 204
Qy	181	KLDLRLDECKASSAKORLKCASLQKFGBRFAKAWAVARLSQRFPAEFAEVSGLVTDLT 240
Db	205	KLEALREKVLSSAKERFKCASLQKFGBRFAKAWAVARLSQRFPAEFAEVSGLVTDLT 264
Qy	241	VTECHGDLLECCADRDADLAKYICENQDSISSKLEKCECEKPELLEKSHCIAEVENDEMPA 300
Db	265	VHKECHGDLLECCADRDADLAKYICENQDSISSKLEKCECEKPELLEKSHCIAEVENDEMPA 324
Qy	301	DLPSLAANDVESKDYCKNYABAKDVLGMFLMFLYEAARHPDYVVLVLLAKTYETLEKC 360
Db	325	DLPSLAANDVESKDYCKNYABAKDVLGMFLMFLYEAARHPDYVVLVLLAKTYETLEKC 384
Qy	361	CAAADPHECYAKVDFEFPKPLVPEEPQNLKQNCLEFQOLGEYKFNALVRYTKKVPQVST 420
Db	385	CATDDPPTCYAKVDFEFPKPLVPEEPQNLKQNCLEFQOLGEYKFNALVRYTKKVPQVST 444
Qy	421	PTLVEVSRNLKGVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCCTES 480
Db	445	PTLVEVSRNLKGVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCCTES 504
Qy	481	LNNRRPCFSALEVDETYVPKEFNAETFTTFHADICTLSEKEROIKKQTA:VELLVKKHKPAT 540
Db	505	LNNRRPCFSALEVDETYVPKEFNAETFTTFHADICTLSEKEROIKKQTA:VELLVKKHKPAT 564

QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKETCHAFEEGKGLVAASQAAL 583
 DB 565 DEQLKTVMGDFGAFVEKCKCAAENKEGCFSEEGPKLVAAAQAAL 607

RESULTS

```

RESOLUTION
ID ALBU_HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
DE ALB.
GN Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OC NCBI_TaxID=9796;
RN [1]
SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC
RC TISSUE=Liver;
RA MEDLINE=93344595; PubMed=8344282;
RX Ho J.X., Hollowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
RT at 0.27-nm resolution.";
RL Eur. J. Biochem. 215:205-212(1993).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

```

```

EMBL; X74045; CAA52194.1; -.
DR DR
DR DR PIR; S34053; ABHOS.
DR DR HSP; P02768; 1E7B.
DR DR
DR DR InterPro; IPR000264; Serum_albumin.
DR DR
DR DR Pfam; PF00273; transport_prot; 3.
DR DR PRINTS; PR00802; SERUMALBUMIN.
DR DR ProDom; PD002486; Serum_albumin; 1.
DR DR SMART; SM00103; ALBUMIN; 3.
DR DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461

```

FT	DISULFID	460	471	
FT	DISULFID	484	500	
FT	DISULFID	499	510	
FT	DISULFID	537	582	
FT	DISULFID	581	590	
SQL	SEQUENCE	607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;		

Query Match
 Best Local Similarity 79.8%; Score 2475.5; DB 1; Length 607;
 Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1

Qy	1	DAKSEVAHRKFDLGEENFKALVLIATAQYLQCCPDEDHVKLNVETFAKTCVADESAE	60
Db	25	DTHKSEIARRFNDLGERHFGVLVAFVSQYLQCCPFEDHVKLNVETFAKKADESAE	84
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNEBCFLQHKDDNPNIPLRVPEV	120
Db	85	NCDKSLHTLFGDKLCTVATLRAATYGELADCCOEPRNEBCFLTHKDDHPNLPKL-KPEP	143
Qy	121	DWMTCTAFHDNSETFLKKYLIARHPHYFYFAPELLFFAKRYKAAFTCCQADKAACLLP	180
Db	144	DAQCAAFQEDPDKFLGKLYLXVARHPHYFYGYPELLFHAEEYKADFTCCPADDKAACLLP	203
Qy	181	KLDELRDGKASSAKQRLKCAQSLQKFGERAFKAWAVARLSQRPPEKFAEVSXLVTDLT	240
Db	204	KLDALKERILLSSAKERLKCSSFQNGFERAVKAWAVARLSQKFPKADFAEVSXLVTDLT	263
Qy	241	VHTECHGDLJECADDRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVEND	300
Db	264	VHKECHGDLJECADDRADLAKYICEHQDSISGKLKACCDKPLQKSHCIAEYKEDDLES	323
Qy	301	DLPSLAADPVESKQVCKNYAEAKQVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLEK	360
Db	324	DLPALAAPDAEDKELCKRYKQAKQVFLGTFLYEVARRHPDYSVSLLRRIAKTYEATLEK	383
Qy	361	CAAADPHCYAKVDFDEPKPLVEEPONLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST	420
Db	384	CAEADPPACYRTVFDQFPLVEEPKSLVKKNCDLFEVGEYDFQNALIVRYTKKAPQVST	443
Qy	421	PTLVEVSRLNGKVGSKCCKHPEAKRMPCAEDYLSVTLNQLCVLHEKTPUSDRVTKCTTES	480
Db	444	PTLVEIGRTLGVKSGRCCKPLPESERLPCSENHALALNRLCVLHEKTPVSEKTIKCTCDS	503
Qy	481	LVNRRPCFSALVDEVTYVPKCFNAETTFHADICTLSEKSRQIKOTALVELVKKHPKAT	540
Db	504	LAERRPCFSALLELDGYVPKCFNAETTFHADICTLPEDEKQIKKQALAEVLKHPKAT	563
Qy	541	KEQLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAAL	583
Db	564	KEQLKTVLGNFSAFVAKCCGREDKEACFAEBGPKLVASSOLAL	606

RESIII.T 6

RESULTS:
ALBU_BOVIN
ID ALBU BOVIN STANDARD; PRT; 607 AA.
AC P02769; Q02787;
DT 21-JUL-1986 (Rel. 01, Created)
DD 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
CS Bos taurus (Bovine) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
RN SEQUENCE FROM N.A.
RP Holowachuk E.W., Stoltenberg J.K., Reed R.G., Peters T. Jr.;
RA Submitted (AUG-1991) to the EMBL/genBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A., AND VARIANT THR-214.
RP TISSUE=Liver;
RC


```

Db 204 K1ETWREKVLASSARQRLRCAS:QKFGERALKAWSVAR:SQKFPKAEFVEVTKLVDTLTK 263
Qy 241 VHTCECHGDLLEACADRADLAKYICENODSISSKLECECKPLLEKSHCIAEVENDEMPA 300
Db 264 VHKCECHGDLLEACADRADLAKYICENODSISSKLECECKPLLEKSHCIAEVENDEMPA 323
Qy 301 DLPSSAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYVSVLLRLAKTYE:TTLEKC 360
Db 324 NLLPPLTADPAEDKDVCKNYAEAKDVLGMFLYFYARRHPDYVSVLLRLAKTYE:TTLEKC 383
Qy 361 CAADPHCEYAKVDFEFLPVEEPONLIKONCELFQELGEYKFNALLVRYTKVPOVST 420
Db 384 CAKDDPHACYSTVDFKLKHLVDPEONLIKONCELDQKELGEYKFNALLVRYTKVPOVST 443
Qy 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVNLCLVHLHEKTPVSDRVTKCCTES 480
Db 444 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVNLCLVHLHEKTPVSDRVTKCCTES 503
Qy 481 LVNRRPCFSALVDEVTYVPKEFNAETFTFHAD:CTLSEKERQ:KKQTALVELVKKPKAT 540
Db 504 LVNRRPCFSALVDEVTYVPKEFNAETFTFHAD:CTLSEKERQ:KKQTALVELVKKPKAT 563
Qy 541 KEQLKAVMDDFAAVFEKCKKADKCTCFABEGKLVAAASQAL 583
Db 564 EQQLKAVMDDFAAVFEKCKKADKCTCFABEGKLVAAASQAL 606

RESULT 7
ALBU RABIT
ID ALBU RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last: sequence update)
DT 28-FEB-2003 (Rel. 41, Last: annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U18344; AAB58347.1; -.
CC HSSP; P02768; 1E78.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24

```

```

FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; CF5B92647AABF9A2 CRC64;

Query Match 78.8%; Score 2446; DB i; Length 608;
Best Local Similarity 74.1%; Pred. No. 8.1e-152;
Matches 433; Conservative 77; Mismatches 74; Indels 0; Gaps 0;

Qy 1 DAHSEVAHREKOLGEENFKALVLIIFAQVLOCPREDHVLYNEVTEPAKTCVADESAA 60
Db 25 EAHKSEIAHRENDVGEHFIGLVLIITFSQYLQKCPVEEHAFLVKEVTDLAAACVADESAA 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKODNPNLRLVREPV 120
Db 85 NCDKSLHDI FGDKLCPSLRDTYGVADCCCKEPERNECFLLHKKDDKDELPPARPEA 144
Qy 121 DVMTAFHNEETPLKYLVEIARRHPYFYAPPELLFPKRYKAAFTCCCAAKAACLLP 180
Db 145 DVLCKAFHDDKAPFFGHLYVEARRHPYFYAPPELLYVYAKYKAILTECCEADKGACLTP 204
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFVSKLVDTLTK 240
Db 205 KLDALEGKSLISAAQERLCASLOKFGDRAYKAWALVRLSQRFPKADFTDISKIIVDTLTK 264
Qy 241 VHTCECHGDLLEACADRADLAKYICENODSISSKLECECKPLLEKSHCIAEVENDEMPA 300
Db 265 VHKCECHGDLLEACADRADLAKYICENODSISSKLECECKPLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSSAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYVSVLLRLAKTYE:TTLEKC 360
Db 325 GLPAVABEFVEDKDVCKNYEAKDLFLGKFLYFYARRHPDYVSVLLRLGKAYEATLKKC 384
Qy 361 CAADPHCEYAKVDFEFLPVEEPONLIKONCELFQELGEYKFNALLVRYTKVPOVST 420
Db 385 CATDDPHACVAKVDFEFLPVEEPONLIKONCELFQELGEYKFNALLVRYTKVPOVST 444
Qy 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVNLCLVHLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVNLCLVHLHEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALVDEVTYVPKEFNAETFTFHAD:CTLSEKERQ:KKQTALVELVKKPKAT 540
Db 505 LVNRRPCFSALVDEVTYVPKEFNAETFTFHAD:CTLSEKERQ:KKQTALVELVKKPKAT 564

ALBU SHEEP
ID ALBU SHEEP STANDARD; PRT; 607 AA.
AC P14639;

```

RESULT 8

ALBU SHEEP

ID ALBU SHEEP

AC P14639;

STANDARD;

PRT;

607 AA.

"Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage of rat serum albumin.",
 J. Biochem. 83:35-48(1978).
 [4]
 SEQUENCE OF 223-288 AND 572-608.
 MEDLINE=76260153; PubMed=956149;
 Isemura S., Ikenaka T.;
 "Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequences of four fragments.",
 J. Biochem. 79:1183-1196(1976).
 [5]
 SEQUENCE OF 166-174.
 TISSUE=Plasma;
 MEDLINE=87194805; PubMed=2437111;
 Caraway R.E., Mitra S.P., Cochran D.E.;
 "Structure of a biologically active neurotensin-related peptide obtained from pepsin-treated albumin(s).",
 J. Biol. Chem. 262:5968-5973(1987).
 [6]
 COPPER-BINDING.
 MEDLINE=79001617; PubMed=80265;
 Aoyagi Y., Ikenaka T., Ichida F.;
 "Copper(II)-binding ability of human alpha-fetoprotein.",
 Cancer Res. 38:3483-3486(1978).
 -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
 -!- FUNCTION: NRP REGULATES FAT DIGESTION, LIPID ABSORPTION, AND BLOOD FLOW (POTENTIAL).
 -!- SUBCELLULAR LOCATION: Secreted.
 -!- TISSUE SPECIFICITY: Plasma.
 -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 -!- SIMILARITY: Contains 3 albumin domains.

 This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; V01222; CAA24532.1; -.
 PIR; A93872; ABR7S.
 HSP; P02768; 1E7B.
 InterPro; IPR000264; Serum_albumin.
 Pfam; PF00273; transport_prot; 3.
 PRINTS; P00802; SERUMALBUMIN.
 ProDom; PD002486; Serum_albumin; 1.
 SMART; SM00103; ALBUMIN; 3.
 PROSITE; PS00212; ALBUMIN; 3.
 Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT SERUM ALBUMIN.
 FT NEUTROTENSIN-RELATED PEPTIDE.
 FT PEPTIDE 166 174
 FT DOMAIN 1 25
 FT ALBUMIN 1.
 FT DOMAIN 212 397
 FT ALBUMIN 2.
 FT DOMAIN 404 595
 FT ALBUMIN 3.
 FT METAL 27 27
 FT COPPER.
 FT METAL 77 86
 FT BY SIMILARITY.
 FT DISULFID 77 86
 FT BY SIMILARITY.
 FT DISULFID 99 115
 FT BY SIMILARITY.
 FT DISULFID 114 125
 FT BY SIMILARITY.
 FT DISULFID 148 193
 FT BY SIMILARITY.
 FT DISULFID 192 201
 FT BY SIMILARITY.
 FT DISULFID 224 270
 FT BY SIMILARITY.
 FT DISULFID 269 277
 FT BY SIMILARITY.
 FT DISULFID 289 303
 FT BY SIMILARITY.
 FT DISULFID 302 313
 FT BY SIMILARITY.
 FT DISULFID 340 385
 FT BY SIMILARITY.
 FT DISULFID 384 393
 FT BY SIMILARITY.
 FT DISULFID 416 462
 FT BY SIMILARITY.

FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 495 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT VARIANT 262 262 V -> L.
 FT CONFLICT 174 174 Y -> L (IN REF. 5).
 SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;
 Query Match 78.2%; Score 2426; DB 1; Length 608;
 Best Local Similarity 73.4%; Pred. No. 1.6e-150;
 Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRFKDLGSENFKALVIAFAQYLOQCQFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 25 EAHKSIAHRPKDLGSEHFKGLVIAFSQYLOKCPVEEHIKLVQEVTFDAKTCVADESAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRTYTGEMACCAKQEPERNECFLOHKDDNPRLVLRPEV 120
 DB 85 NCDKSIHTLFGDKLCAIPKLRDNYGELADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144
 QY 121 DVMCTAFHNEETFLAKLYELARRHPYFAPPELLFAFKRYKAAFTCCCAADKRACLLP 180
 DB 145 EAMCTSFQENPTSFZGHYLHEVARRHPYFAPPELLYAEKYNEVLTCCTESDAACLTP 204
 QY 181 KLDELDEGKASSAKQRLKASLQKFEGERAFKAWAVARLSQRPFPKAEFAEVSKLVTDLT 240
 DB 205 KLDVKEKALVAVRQRMKSSMQREFGERAFKAWAVARMSQRPFPNAEFAEITKLADVT 264
 QY 241 VHTECCHGOLLECCADDDRADLAKYICENQDSISSKLKECKEKLPLEKSHCIAEVENDE 300
 DB 265 INKECHGDLLECCADDDRAELAKYICENQATISSKLQACCDKPVLOKSQLAEHNDIPA 324
 QY 301 DLPSLAADFVESKDVCKVAEAKDVLGMLFVEYARRHPDYSVLLRLAKTYETTELK 360
 DB 325 DLPSLAADFVEDKEVCNKAABKDFLGLFVYBRRHPDYSVSLRLAKTYEATLEK 384
 QY 361 CAAADPHECYAKVDFEKLVEEPQNLIKQNCLEFQGLGEYKFNQALLVRYTKKVPQVST 420
 DB 385 CAEGDPPACVGTVAEFPQPLVEEPKLVKTNCELYEKLGEYFQNAVLYRTQKAPQVST 444
 QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVNLQCLVHEKTPVSDRVTKCTES 480
 DB 445 PTLVEAARNLGRVTKCTCLPEAQRLPCVEDYLSAILNRLCVJHEKTPVSEKVTCKCSG 504
 QY 481 LVNRRPCFSALEVDVTPKPEAFNAETFTTHADICTLSEKEROIKKOTALVELVKHKPKAT 540
 DB 505 LVERRPCFSALTVDVTPKPEFAETFTPHSDICTLPDKKQIKKOTALAEJVKHKPKAT 564
 QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKLVAAQAAAL 583
 DB 565 EDQLKTVMGDFAQFVCKCKKADKDNCFATGPNLVARSKEAL 607
 RESULT 10
 ID ALBU_PIG STANDARD; PRT; 605 AA.
 AC P08835; O29018;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (Fragment).
 GN ALB.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA MEDLINE=89016582; PubMed=3174440;
 RA Baldwin G.S., Weinstock J.;
 RT "Nucleotide sequence of porcine liver albumin.";

DR PRINTS: PRO0802; SERUMALBUMIN.
 DR PRODOM; P003486; Serum albumin; 1.
 DR SMART; SVO0103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 609 SERC ALBUMIN.
 FT DOMAIN 25 206 ALBUMIN 1.
 FT DOMAIN 213 398 ALBUMIN 2.
 FT DOMAIN 405 596 ALBUMIN 3.
 FT METAL 28 28 COPPER.
 FT DISULFID 78 87 BY SIMILARITY.
 FT DISULFID 100 116 BY SIMILARITY.
 FT DISULFID 115 126 BY SIMILARITY.
 FT DISULFID 149 194 BY SIMILARITY.
 FT DISULFID 193 202 BY SIMILARITY.
 FT DISULFID 225 271 BY SIMILARITY.
 FT DISULFID 270 278 BY SIMILARITY.
 FT DISULFID 290 304 BY SIMILARITY.
 FT DISULFID 303 314 BY SIMILARITY.
 FT DISULFID 341 386 BY SIMILARITY.
 FT DISULFID 385 394 BY SIMILARITY.
 FT DISULFID 417 463 BY SIMILARITY.
 FT DISULFID 462 473 BY SIMILARITY.
 FT DISULFID 486 502 BY SIMILARITY.
 FT DISULFID 501 512 BY SIMILARITY.
 FT DISULFID 539 584 BY SIMILARITY.
 FT DISULFID 583 592 BY SIMILARITY.
 SQ SEQUENCE 609 AA; 68940 MW; 9CASF97F67EF1A8 CRC64;
 Query Match 76.9%; Score 2387; DB 1; Length 609;
 Best Local Similarity 73.9%; Pred. No. 5.5e-148;
 Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;

QY 2 AHKSVARFDFKLGKFNKALVLIIFAQYLOQCFEDHVKLVNTEPAKTCVADESSEN 61
 DB 27 AHKSIHRYKDLGKFKYFGVLYTFYQYLOKCSYEEHVKLVREITDFASCAKDESSEN 86
 QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVVRPEVD 121
 DB 87 CDKSLHTLFGDKLCSYFNGEKYVADMACCAKQEPERNECFLOHKDDNPNLPRVVRPEVD 146
 QY 122 VNCATPHDNEETFLKXYIETARRHPYVYAPPELLFAXRYKAAFTCCOAAADKAAACLLPK 181
 DB 147 AMCTAFQENAEAFMGHYLHEVARRHPYFYGPPELLYADKYTAVLTCCCAADKGAACLLPK 206
 QY 182 LDELEDEGKASSAKOELKASLOKTERAFKAWARLSORPPKAEFAEVSKVLTDLTKV 241
 DB 207 LDALKEKALVSAVRORLKCSSMKKFGERAFKAWARMSQTFPNADFABITKLAIDLTKV 266
 QY 242 HTECCGGLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPAD 301
 DB 267 TQECCHGGLLECADRAELAKYMCNQNQASISSKLQACCCKENLQSQSLAEVENDEMPAD 326
 QY 302 LPSLAADPVESKDCVXAEAKDVF.LGMFLYEVARRHPDYSVLLRLRAKTYETTELKCC 361
 DB 327 LPALTADTFVEDKDCVXAEAKDVF.LGFLYEVARRHPDYSVLLRLRAKTYETTELKCC 386
 QY 362 AAADHECVAKYFDFEKLPLVEEPQNLKONCELFQOLGEYFQNALVRYTKYQVNSTP 421
 DB 387 AEAHPHACVGHVDFEKLPLVEEPQNLKONCELYEKLGEYFQNALVRYTKYQVNSTP 446
 QY 422 TLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQCVLHEKTPVSDRVTKCTESL 481
 DB 447 TLVEAARSIGRVTHCCALPEKRLPCVEDYLSALINRVCLLHEKTPVSEQVTKCCSSL 506
 QY 482 VNRPCFSALEVDETVYVPKFEAFNFTTFHADICTLSEKERQIKKOTALVELVFKHKKPKATK 541
 DB 507 VERRPCFSALEVDETVYVPKFEAFNFTTFHADICTLSEKERQIKKOTALVELVFKHKKPKATE 566
 QY 542 EQLKAWMDFAFVKECKKADDKETCTFAEEGKKLVAASQAAL 583

DB 567 EQLKVMGDFABEFLEKCKQEDKEACFSTEGPKLVAESQAL 608
 RESULT 12
 ALBU_MOUSE
 ID ALBU_MOUSE STANDARD; PRT; 608 AA.
 AC P07724; Q61802;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB OR ALB1 OR ALB-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA F. Reichmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staibli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Morbaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Hayashizaki Y.;
 RL "Functional annotation of a full-length mouse cDNA collection.";
 RN Nature 409:685-690(2001).
 RP SEQUENCE OF 99-516 FROM N.A.
 RX MEDLINE=88216123; PubMed=2452956;
 RA Xinghetti P.P., Law S.W., Dugaiczky A.;
 RT "The rate of molecular evolution of alpha-fetoprotein approaches that
 of pseudogenes.";
 RL Mol. Biol. Evol. 2:347-358(1985).
 RN [4]
 RP SEQUENCE OF 477-551 FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE=90269606; PubMed=1971802;
 RA Boccaccio C., Deschatrette J., Meunier-Rotival M.;
 RT "Empty and occupied insertion site of the truncated LINE-1 repeat
 located in the mouse serum albumin-encoding gene.";
 RL Gene 88:181-186(1990).
 RN [5]
 RP SEQUENCE OF 25-44.
 RC TISSUE=Liver;
 RX MEDLINE=93162044; PubMed=1286668;
 RA Giometti C.S., Taylor J., Tollaksen S.L.;
 RT "Mouse liver protein database: a catalog of proteins detected by two-
 dimensional gel electrophoresis.";
 RL Electrophoresis 13:970-991(1992).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X60688; CAA43098.1; -;
 DR EMBL; V00381; CAA23680.1; -;
 DR PIR; S15573; ABCHS;
 DR HSP; P02768; IE7B;
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN_3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 23
 FT CHAIN 24 615 SERUM ALBUMIN.
 FT DOMAIN 24 209 ALBUMIN 1.
 FT DOMAIN 216 401 ALBUMIN 2.
 FT DOMAIN 408 599 ALBUMIN 3.
 FT METAL 30 30 COPPER (BY SIMILARITY).
 FT DISULFID 80 89 BY SIMILARITY.
 FT DISULFID 102 118 BY SIMILARITY.
 FT DISULFID 117 128 BY SIMILARITY.
 FT DISULFID 152 197 BY SIMILARITY.
 FT DISULFID 196 265 BY SIMILARITY.
 FT DISULFID 228 274 BY SIMILARITY.
 FT DISULFID 273 281 BY SIMILARITY.
 FT DISULFID 293 307 BY SIMILARITY.
 FT DISULFID 306 317 BY SIMILARITY.
 FT DISULFID 344 389 BY SIMILARITY.
 FT DISULFID 398 397 BY SIMILARITY.
 FT DISULFID 420 466 BY SIMILARITY.
 FT DISULFID 465 476 BY SIMILARITY.
 FT DISULFID 489 505 BY SIMILARITY.
 FT DISULFID 504 515 BY SIMILARITY.
 FT DISULFID 542 587 BY SIMILARITY.
 FT DISULFID 586 595 BY SIMILARITY.
 FT CARBOHYD 500 500 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 24 24 F -> M (IN REF. 3).
 SQ SEQUENCE 615 AA; 69918 MW; E59E4B5CAEC066C6 CRC64;

Query Match 50.2%; Score 1557.5; DS 1; Length 615;
 Best Local Similarity 46.7%; Pred. No. 4,3e-94;
 Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;

QY 3 HKSEVAHFKDGEENFKALVLAFAQYLOQCFEDHVKLVNEVTEFAKTCVADESAENC 62
 DB 30 HKSEIAHRYNDLKEETFKAVAMITFAQYLRQCSYEGSLKLVNDVVDLAQKCVANEADAEC 89
 QY 63 DKSLHTLFGDKCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNNLPR-LVRPEVD 121
 DB 90 SKPLPSIILDETQCEVLEKRDSDVGAMADCCSKADPERNECFJ5FKVSDPFDVQYQRRASD 149
 QY 122 VMCTAFHDNEETFLKYLVEIARRHRYFYAPPELLFFAKRYKAAFTCCQAADKAACLLPK 181
 DB 150 VICQEQDNRVSLFGLHFIYSVARRHFFLYAPILSPAVDFEHALQSCCKESDVGACLTK 209
 QY 182 LDELDEGKASSAKQLKASLOKGERAFKAWAVARLSQRFPKAPAEVSKLVTLTKV 241
 DB 210 EIVMREKAGVSVKQVFCGLIKQFGRVFRQRLTYLSOKYPKAPFSEVSKFVHDSIGV 269
 QY 242 HTECHGDLLECCADPADLAUKYICENQDSISSKLKCECKPILKSHCIARVENDEMDAD 301
 DB 270 HKECEGDMVECDMDMARMMNLCSQQDFVSGKIKDCCCKPIVERSQCIEMAEAFDEKPAD 329
 QY 302 LPSLAADFVSKDCKVNAEAKDVF-GMFLXYARRHDPYSVVLLRLAKTYETTLKCC 361
 DB 330 LPSLVEKYIEDKVCVSKFEAGHDAFMAEFVYYSRRHPFESQLINRIAKGVESLJECCK 389

QY 362 AAADPHCYAKVDFDEPKPLVEBPQMLIKONCELFEOLOGYKFQNALLVRYTKVPOVSTP 421
 DB 390 KTDNPAECYANQEQUNQHIKETQDVVTKNCCOLLHGHGEADFLKSLILRYTKMPOVPTD 449
 QY 422 TLVVEGRNLGVSGKCKRHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCCTESL 481
 DB 450 LLLCTGKMTTIGTKCCQLGEDRRMACSGYLSIVIHDTCRKOETTPINDNVSCCSQJY 509
 QY 482 VNERPCFSALEVDETYVPKEFAETFTFHADICTLSEKERQIKKOTALVELVKKHEKATK 541
 DB 510 ANRRPCFTAMGVDTKYVPPFPDMFSDFKLCSAPAESEVGGQMKLLINLKRKPQMTF 569
 QY 542 EQLKAVMDDFAAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
 DB 570 EQIKTIADGFTAMVDKCKCKQSDINTCFEGEGANLIVQSRATLGI 613

RESULT 14
 PETA_PANTR
 ID PETA_PANTR STANDARD; PRT; 609 AA.
 AC Q28789;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alpha-fetoprotein precursor (Alpha-fetog-obulin) (Alpha-1-fetoprotein).
 DE fetoprotein).
 GN AFP.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96032345; PubMed=7557431;
 RA Nishio H., Gibbs P.E., Minghetti P.P., Zielinski R., Dugaiczky A.;
 RT "The chimpanzee alpha-fetoprotein-encoding gene shows structural
 RT similarity to that of gorilla but distinct differences from that of
 RT human.";
 RL Gene 162:1213-220(1995).
 CC -!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND
 CC BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE
 CC (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.
 CC -!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION
 CC TO THE MONOMERIC FORM.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and
 CC yolk sac.
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U21916; AAA91641.1; -;
 CC PIR; JC4258; JC4258.
 CC HSP; P02768; IE7B.
 CC InterPro; IPR000264; Serum_albumin.
 CC Pfam; PF00273; transport_prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum_albumin; 1.
 CC SMART; SM00103; ALBUMIN_3.
 CC PROSITE; PS00212; ALBUMIN; 2.
 KW Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
 KW Signal.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 609 ALPHA-FETOPROTEIN.
 FT DOMAIN 20 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.

```

FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 22 22 COPPER AND NICKEL (BY SIMILARITY);
FT DISULFID 99 114 BY SIMILARITY.
FT DISULFID 113 124 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL);
FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENT-AL);
SQ SEQUENCE 609 AA; 68741 MW; C032987CADCE672B CRC64;

Query Match 40.4%; Score 1253.5; DB 1; Length 609;
Best Local Similarity 40.1%; Pred. No. 2.4e-74;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY 3 HKSE-----VAHRRKDLGEEFNFKALVLIAPAOYLQOCFFEDHVKLVNEVTEFAKTCVADE 57
DB 22 HNEVGIASILDYSOCTAEINLDTATIFFAQFVGEATYKSKRMVKDALTAEKPTGDE 81
QY 58 SAENCDKSJHTLFGKLCQVATLRETYGEMADCCAKQEPERNECFLOHQKDNPP-NLPLRV 116
DB 82 QSAGCLENLPAFLBELCREKLEKYGH-SDCCSQSEGRHNCFLAHKKPTPASIFFQ 140
QY 117 REVDMCTAFDNDEETLKKVLYBIARRHPYFYAPELLFFAKRYKAFTSCCAAKAA 176
DB 141 VPEPTSCAEYEDRETFMKNFIYIARRHPFLYAPTILLWAARYDKTIPSCCAENAVE 200
QY 177 CLLPKLDLRDEGKASSAKQRLKUCASLOKFGERAFKAWAVARLSORPPKFAFVSKLVT 236
DB 201 CQTQKAATVKELRESSLLNQACAMKNFGTRTFOA-TVTKLSQKFTKVNFTHIQKLVL 260
QY 237 DLTKVHTCCGGDLLECADDRADLAKYICENQDISSKLKECEKPLLEKSHCIAEVND 296
DB 261 DVARVHECHCCGVDLQDQGEKIMSYCSQODT-SNKITECKLTTLERQCCIHAEND 320
QY 297 ENPADLPSLADPVESKQVKNYAKQVFLQMGFLMELYEYARRHPDYSVVLRLAKTYETT 356
DB 321 EKPEGLSPNLNFGLDRDFNCFSSGGEKNIFLASFVHEYSRRHPQJAVSVILRVAKGYOEL 380
QY 357 LEKCCAAADPHCYAKVDEPFKPLVDEEPQNLIKQNCLEFEOLOGYKFNALLVRYTKVP 416
DB 381 LEKCFQTEPLECDQGEELQKYIOESQALAKRSCGJFQKLGELYLQNALVAYTKAP 440
QY 417 QVSTPTLVEVSNLQVSKCKKHPEAKRMPCAEDYLSVLNLQCVLHEKTPVSDRVTKC 476
DB 441 QLSSELMAITKMAATAATCQLQSEDKLLACGEAADIIGHLCIRHETTPVNPVGVC 500
QY 477 CTESLVNRPCPSALEVEDYVPKFFNAETFTFHADICTLSEKERQIKKCTALVELVKHK 536
DB 501 CTSSVANRRPCFSSLVVDYTPPAFSDOKFLFHKDLCCAGVALQTKQBFNLNVKQK 560
QY 537 PRATKEQLKAVNDDFAAFVEKCKQADDDCTCFABEGKKLVAAASCAALGL 585
DB 561 PQITEOLEAVTADFSGLLEKCCQGEQEVCFABEGSKLISKTRALGV 609

```

RESULT 15

FETA HUMAN

ID FETA HUMAN

AC P02771; STANDARD; PRT; 609 AA.

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).

DE APP.

GN Homo sapiens (Human).

CS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

OX [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=83273664; PubMed=6192439;

RX Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;

RA "Primary structures of human alpha-fetoprotein and its mRNA.;"

RT Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608(1983).

RL [2]

RN SEQUENCE FROM N.A.

RP MEDLINE=87185438; PubMed=2436661;

RX Gibbs P.E.M., Zielinski R., Boyd C., Dugaiczky A.;

RA "Structure, polymorphism, and novel repeated DNA elements revealed by a complete sequence of the human alpha-fetoprotein gene.;"

RT Biochemistry 26:1332-1343(1987).

RL [3]

RN SEQUENCE OF 1-28 FROM N.A.

RP MEDLINE=93278385; PubMed=7684942;

RX McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M., Tilghman S., Krumlauf R., Tuddenham E.G.D.;

RA "A G->A substitution in an HNF I binding site in the human alpha-fetoprotein gene is associated with hereditary persistence of alpha-fetoprotein (HPAFP).;"

RT Hum. Mol. Genet. 2:379-379(1993).

RL [4]

RN SEQUENCE OF 429-556 FROM N.A.

RP MEDLINE=83158778; PubMed=6187626;

RX Beattie W.G., Dugaiczky A.;

RA "Structure and evolution of human alpha-fetoprotein deduced from partial sequence of cloned cDNA.;"

RT Gene 20:415-422(1982).

RL [5]

RN PARTIAL SEQUENCE OF 19-609.

RP MEDLINE=91242409; PubMed=1709810;

RX Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F., Ceccarini C., Terrana B.;

RA "Human alpha-fetoprotein primary structure: a mass spectrometric study.;"

RT Biochemistry 30:5061-5066(1991).

RL [6]

RN PRELIMINARY SEQUENCE OF 19-35.

RP MEDLINE=77242506; PubMed=70228;

RX Yachnin S., Hsu R., Heinrikson R.L., Miller J.B.;

RA "Studies on human alpha-fetoprotein. Isolation and characterization of monomeric and polymeric forms and amino-terminal sequence analysis.;"

RT Biochim. Biophys. Acta 493:418-428(1977).

RL [7]

RN PRELIMINARY SEQUENCE OF 19-38.

RP MEDLINE=78001760; PubMed=711198;

RX Aoyagi Y., Ikenaka T., Ichida F.;

RA "Comparative chemical structures of human alpha-fetoproteins from fetal serum and from ascites fluid of a patient with hepatoma.;"

RT Cancer Res. 37:3663-3667(1977).

RL [8]

RN PRELIMINARY SEQUENCE OF 19-39.

RP MEDLINE=75018719; PubMed=4138095;

RX Ruoslahti E., Pihko H., Vaheri A., Seppala M., Virolainen M., Kontinen A.;

RA "Alpha fetoprotein: structure and expression in man and inbred mouse strains under normal conditions and liver injury.;"

RT Johns Hopkins Med. J. Suppl. 3:249-255(1974).

RL [9]

RN GENE STRUCTURE

RP MEDLINE=85182629; PubMed=2580830;

RX Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G., Tamaoki T.;

RA "The human alpha-fetoprotein gene. Sequence organization and the 5'

Search completed: October 27, 2003, 15:05:17
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:01:33 ; Search time 105 Seconds

(without alignments)

1437.722 Million cell updates/sec

Title: US-09-832-501-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHFRLDGEENFK.....TCFAEFGKLVAAASQAALGL 565

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 259052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMS_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvitus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	80.7	608	5	Q95VB7
2	2374	76.5	608	11	Q8C7H3
3	2336	75.3	576	11	Q8C7C7
4	1865.5	60.1	396	4	Q8IUK7
5	1295.5	41.7	527	13	Q8JIA9
6	1242	40.0	609	6	Q8M3U5
7	1242	40.0	626	13	Q8UM05
8	1218.5	39.3	610	6	Q8MJ76
9	1087	35.0	624	13	Q8UM06
10	1082	34.9	605	11	Q8BK65
11	1076	34.7	605	11	Q8BK56
12	1045	33.7	400	13	Q8JIA7
13	955	30.8	603	13	Q9YGH6
14	928.5	29.9	614	13	Q9L134
15	888	28.6	406	13	Q8JIA8
16	739	23.8	205	11	Q8C7H4

17	723	23.0	417	11	Q8ROJ9
18	376.5	12.1	484	13	Q9W6F5
19	373	12.0	476	11	Q9CY31
20	372	12.0	476	11	Q9IXG1
21	365	11.8	551	13	O42279
22	331	10.7	122	13	Q90WZ8
23	290	9.3	123	13	Q90WZ6
24	264	8.5	135	11	Q63205
25	188	6.1	1723	2	Q9JMX8
26	184	5.9	1819	16	Q9ZLV0
27	184	5.9	1927	16	O25262
28	162.5	5.2	44	6	Q95WC2
29	162.5	5.2	680	5	Q9V6S8
30	161	5.2	1079	3	Q96V11
31	156	5.0	1026	3	Q74669
32	154.5	5.0	3843	5	Q9USD0
33	153.5	4.9	661	5	Q8MS79
34	153.5	4.9	3843	5	Q9VU94
35	153	4.9	62	6	Q8M1L1
36	150	4.8	1065	3	Q01828
37	149	4.8	40	6	Q9TRAS
38	148	4.8	1028	3	O74668
39	144.5	4.7	8749	4	Q8NF91
40	142.5	4.6	2755	10	Q9LJ70
41	141.5	4.6	1560	5	Q26644
42	138.5	4.5	1069	3	Q96V12
43	137.5	4.4	1348	16	Q8YK55
44	135.5	4.4	2841	5	Q8MLU9
45	135.5	4.4	2931	5	Q9W2C6

ALIGNMENTS

RESULT 1

Q95VB7	PRELIMINARY;	PRT;	608 AA.
AC	Q95VB7;		
DT	01-DEC-2001 (TREMBlrel. 19, Created)		
DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)		
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)		
DE	Albumin.		
OS	Schistosoma mansoni (Blood fluke).		
OC	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;		
OC	Schistosomatidae; Schistosomatidae; Schistosoma.		
OX	NCBI_TaxID=6183;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;		
RT	"Albumin precursor homolog is a novel T helper cell immunogenic egg component in murine infection with Schistosoma mansoni.";		
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF418550; AAL08579.1;		
DR	InterPro; IPR000264; Serum_albumin.		
DR	Pfam; PF00273; transport_prot; 3.		
DR	PRINTS; PR00802; SERUMALBUMIN.		
DR	ProDom; PD002486; Serum_albumin; 1.		
DR	SMART; SM00103; ALBUMIN; 3.		
DR	PROSITE; PS00212; ALBUMIN; 2.		
DR	SEQUENCE 638 AA; 68225 MW; E5EABB28E1C66B54 CRC64;		
Query Match	80.7%;	Score 2504;	DB 5;
Best Local Similarity	76.3%;	Pred. No. 1.7e-198;	
Matches 445;	Conservative 79;	Mismatches 59;	Indels 0; Gaps 0;

Qy	1	DAHKSEVAHFRLDGEENFKLVIAFAOYLCQCPEDHVKLVNVEFATKCVADSEAE	60
Db	25	DAHKSEIAHFRLDGEQHFGLVIAFSPQLQCPVEHVKLVNVEFATKCVADSEAE	84
Qy	61	NCDKSLHTLFDGKLCVTATRETYGEMADCCAKOEPERNECFLOHKDDNPNRLVRPEV	120
Db	85	NCDKSLHTLFDGKLCVATPTLRDLSYGLADCCAKOEPERNECFLOHKDDHNPFPVRPDA	144

```

QY 122 DVMCTAFHNEBTFLKKYLYEIARRHPYFAPPELLFFPAKRYKAAPTECCQAAQKAACLJP 180
DB 145 EAMCTSFQENAVTFMGHLYHEVARHPYFAPPELLYYAEKYSALNTECCGGAACACITP 204
QY 181 KLDELDECKASSAKORLKCASLOKFGGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 205 KLDAUKKALASSVNRLLKCSLQRFQORAFKAWAVARLSQFPKADFAEITKLA-DLTK 264
QY 241 WHITECHGDLLECADDRADLAKYICENODSISSKLECEKPLLEKSHCIAEVENDEMPA 300
DB 265 LTBEECHGDLLECADDRADLAKYICENODSISSKLECEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAAAFVESKOVCKYAEAKDVLGMFLYFYARRHPDYSVLLLRJAKTYETLEKC 360
DB 325 DLPSLAAAFVEDKEVCKYAEAKDVLGMFLYFYARRHPDYSVLLLRJAKTYETLEKC 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKONCELFQGLGEYKFNALLVRYTKVQVUST 420
DB 385 CAADPHCEYAKVDFEFPKPLVEEPONLIKONCELFQGLGEYKFNALLVRYTKVQVUST 444
QY 421 PTLVEVSRNLGVKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540
DB 505 LVNRRPCFSALVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 564
QY 541 KEOLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAAL 583
DB 565 GEOLATVLGEFTAFUDKCKCKEACFSEDGPKLVASSQAAL 607

```

```

RESULT 2
Q8C7H3 PRELIMINARY; PRT; 608 AA.
ID Q8C7H3;
AC Q8C7H3;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Albumin 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
SQ SEQUENCE 608 AA; 68722 MW; 292F60EEDD3A61B4 CRC64;

```

```

Query Match 76.5%; Score 2374; DB 11; Length 608;
Best Local Similarity 72.2%; Pred. No. 2.7e-178;
Matches 421; Conservative 80; Mismatches 82; Indels 0; Gaps 0;

QY 1 DAHSEVARRFKDLGSENFKALVIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 EAHKSIAHYNDLGBQHFGLVLIASFQYLKCSYDEHAKLVQEVTFDAKTCVADESAA 84
QY 61 NCDKSLHTLFGKLCVTATLRETYGEMADCCAKQBPNERNECFLOHKDNDPNLRLVRREV 120
DB 85 NCDKSLHTLFGKLCVATLRETYGEMADCCAKQBPNERNECFLOHKDNDPNLRLVRREV 144
QY 121 DVMCTAFHNEBTFLKKYLYEIARRHPYFAPPELLFFPAKRYKAAPTECCQAAQKAACLJP 180
DB 145 EAMCTSFQENAVTFMGHLYHEVARHPYFAPPELLYYAEKYSALNTECCGGAACACITP 204

```

```

QY 181 KLDELDECKASSAKORLKCASLOKFGGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 205 KLDAUKKALASSVNRLLKCSLQRFQORAFKAWAVARLSQFPKADFAEITKLA-DLTK 264
QY 241 WHITECHGDLLECADDRADLAKYICENODSISSKLECEKPLLEKSHCIAEVENDEMPA 300
DB 265 VNBECCHGDLLECADDRADLAKYICENODSISSKLECEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAAAFVESKOVCKYAEAKDVLGMFLYFYARRHPDYSVLLLRJAKTYETLEKC 360
DB 325 DLPSLAAAFVEDKEVCKYAEAKDVLGMFLYFYARRHPDYSVLLLRJAKTYETLEKC 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKONCELFQGLGEYKFNALLVRYTKVQVUST 420
DB 385 CAADPHCEYAKVDFEFPKPLVEEPONLIKONCELFQGLGEYKFNALLVRYTKVQVUST 444
QY 421 PTLVEVSRNLGVKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540
DB 505 LVNRRPCFSALVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 564
QY 541 KEOLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAAL 583
DB 565 GEOLATVLGEFTAFUDKCKCKEACFSEDGPKLVTRCKDTL 607

```

```

RESULT 3
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7;
AC Q8C7C7;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Albumin 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050644; BAC34360.1; -.
FT NON_TER
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;

```

```

Query Match 75.3%; Score 2336; DB 11; Length 576;
Best Local Similarity 72.2%; Pred. No. 2.5e-175;
Matches 415; Conservative 79; Mismatches 81; Indels 0; Gaps 0;

QY 9 HRFKDLGSENFKALVIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 68
DB 11 NRYNDLGEQHFGLVLIASFQYLKCSYDEHAKLVQEVTFDAKTCVADESAAE 60
QY 69 LFGDKLCVTATLRETYGEMADCCAKQBPNERNECFLOHKDNDPNLRLVRREVDMCTAFH 128
DB 61 LFGDKLCVATLRETYGEMADCCAKQBPNERNECFLOHKDNDPNLRLVRREVDMCTAFH 120
QY 129 DNEBTFLKKYLYEIARRHPYFAPPELLFFPAKRYKAAPTECCQAAQKAACLJP 180
DB 121 ENPTTFMGHLYHEVARHPYFAPPELLYYAEQYNEILLTQCCAEADKESCLTP 180
QY 189 GKASSAKORLKCASLOKFGGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTKVHTECHG 248
DB 181 ALVSSVRQRMCKSSMQKFGGERAFKAWAVARLSOTFPNADFAEITKLA-DLTKVHTECHG 240

```


QY 249 DLLECADDRADLAKYICENODSISSKLKECCCEKPLLEKSHCIAEVENDEMPADLPISLAAD 308
 DB 241 DLLECADDRADLAKYICENODSISSKLKECCCEKPLLEKSHCIAEVENDEMPADLPISLAAD 300
 QY 309 FVSKGVCKVYAEAKDVLGMLFLEYARRPDYSVLLLRKAKTYETTLKCCAAADPHE 368
 DB 301 FVEDQEVCKVYAEAKDVLGMLFLEYARRPDYSVLLLRKAKTYETTLKCCAAADPHE 360
 QY 369 CYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKVPOVSTPTLVEVS 428
 DB 361 CYGTVLAEFQPLVEEPONLIKONCELFQELGEYKFNALLVRYTKVPOVSTPTLVEVS 420
 QY 429 NLGKVGKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCF 488
 DB 421 NLGRVGTCKCTPLPDORLPCVEDYLSAIVNRVLJLHEKTPVSHVTKCCSGSLVRRPCF 480
 QY 489 SALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKATKEQLKAV 548
 DB 481 SALTVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKATKEQLKAV 540
 QY 549 DDFAAVKECKKADDKETCFABEGKLVAAASQAAL 583
 DB 541 DDFAAVKECKKADDKETCFABEGKLVAAASQAAL 575

RESULT 4
 QBIUK7
 ID QBIUK7 PRELIMINARY; PRT; 396 AA.
 AC QBIUK7;
 DT 01-MAR-2003 (TremBLrel. 23, Created)
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
 DE Similar to serum albumin precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBC databases.
 DR EMBL; BC035969; AAH35969.1;
 SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 60.1%; Score 1865.5; DS 4; Length 396;
 Best Local Similarity 63.6%; Pred. No. 1.6e-138;
 Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;
 QY 1 DAHSEVARRFKDLGEENFKALVIAFAQYLQOCPPFEDHVKLVNEVTEFAKTCVADES 60
 DB 25 DAHSEVARRFKDLGEENFKALVIAFAQYLQOCPPFEDHVKLVNEVTEFAKTCVADES 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPLRLVRPEV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPLRLVRPEV 144
 QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 153
 QY 181 KLDELDEGKASSAKORLKASLQKFGFRAPKAWAVARLSORFPKAEFAEVSKLVTDLT 240
 DB 164 ----- 163
 QY 241 VITECCHGDLLECADRADLAKYICENODSISSKLKECCCEKPLLEKSHCIAEVENDEMP 300
 DB 164 ----- 163
 QY 301 DLPSLAADPVESKDVCKVYAEAKDVLGMLFLEYARRPDYSVLLLRKAKTYETTLK 360
 DB 164 ----- 171

QY 361 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKVPOVST 420
 DB 172 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKVPOVST 23;
 QY 421 PTLVEVSRLNGLKVGKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 232 PTLVEVSRLNGLKVGKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 291
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 540
 DB 292 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 35;
 QY 541 KEQLKAVMDDFAAVKECKKADDKETCFABEGKLVAAASQAALGL 585
 DB 352 KEQLKAVMDDFAAVKECKKADDKETCFABEGKLVAAASQAALGL 396

RESULT 5
 QBIUK7
 ID QBIUK7 PRELIMINARY; PRT; 527 AA.
 AC QBIUK7;
 DT 01-OCT-2002 (TremBLrel. 22, Created)
 DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
 DE Serum albumin (Fragment)
 OS Sphenodon punctatus (Hatteria) (Tuatara).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
 OX NCBI_TaxID=8508;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Metcalf V.J.; Brennan S.O.; George P.M.; Charoers G.K.;
 RT "Partial coding sequence for Sphenodon punctatus 68 kDa albumin";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBC databases.
 DR EMBL; AF375971; AAH46104.1;
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport prot.; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 FT NON TER 1
 SQ SEQUENCE 527 AA; 59711 MW; C62B799E387F5929 CRC64;

Query Match 41.7%; Score 1295.5; DB 13; Length 527;
 Best Local Similarity 45.9%; Pred. No. 1.4e-93;
 Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;
 QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPLRLVRPEV 120
 DB 5 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPLRLVRPEV 63
 QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 64 EVSKLYODDLTLGNYIYEVARRHPYLVQPPVFATASLYDEALTKCCTADAKTCFHP 123
 QY 181 KLDELDEGKASSAKORLKASLQKFGFRAPKAWAVARLSORFPKAEFAEVSKLVTDLT 240
 DB 124 RIPPLIEVLKMSNGIQENTCGILKKFGERTLKATKLQMSQKPKADFATINKLVEDITH 183
 QY 241 VITECCHGDLLECADRADLAKYICENODSISSKLKECCCEKPLLEKSHCIAEVENDEMP 300
 DB 184 MHTECCRGDTLLECLDRREALTEVTCSHKDAISSKLPTCCCEKSVLGERCEVRLENDOKPA 243
 QY 301 DLPSLAADPVESKDVCKVYAEAKDVLGMLFLEYARRPDYSVLLLRKAKTYETTLK 360
 DB 244 DLSEIAEYIIDPHVCHOHLAKBOAFKAKFLYEVSRHPPELSTLGLVGGKGVQELLERC 303
 QY 361 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKVPOVST 420
 DB 304 CKTDNPPCYQAADLKKHIAQFQELVQNCNLYNTLGGYLFHALLIRYTKRMPOLTS 363

```
QY 421 PTLVEVSRNLGKVGKCKHPEAKRMEPCADYLSVNLCLCVLHEKTPVSDRVTCKCTESLVNRR3PFCFSALEVC 494
Db 364 BELFYTR-ITKAASRCCEVSVDKLCFTSGYVDFVLGQICQHRQSSINNVVQCSCNS 422
QY 481 LVNRRPCFSALEVDYVYPKEFNAETFTPHAD:CTLSEKERQIKKQTAALVELVGHKPKATKEOLKAVMDDFAA 540
Db 423 YALRSLCITS:GGDEKFPVPIEFSADLFTFHEDLCHAACDKLQERKQMIUNLVKHKPNIT 482
QY 541 KEQLKAVMDDFAAVEVCCKADDETCFAEBEGKLVAAASQAL 583
Db 483 KEQLQTVFGFTQMTXCKCAEDHEADCEGEGPKLVAAESQAL 525

RESULT 6
Q8MJUS PRELIMINARY; PRT: 609 AA.
AC Q8MJUS;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Alpha-fetoprotein.
GN AFP.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Furuichi M., Neo S., Hiasue M., Tsuchiya R., Watanabe M.,
RA Hashizaki K., Hisamatsu S., Yamada T.;
RT "Canine alpha-fetoprotein cDNA."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089789; BAC07513.1; -
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN_3.
DR PROSITE; PS00212; ALBUMIN_3.
SQ SEQUENCE 609 AA; 68782 MW; B54B8250C5AF2AF3 CRC64;

Query Match 40.0%; Score 1242; DB 6; Length 609;
Best Local Similarity 40.5%; Pred. No. 2.7e-89;
Matches 231; Conservative 113; Mismatches 225; Indels 2; Gaps 2;

QY 16 BENFKALVLIAPAYLOCCPDEHVKLVNEVTEFAKTCVADESAENCDKS:HTLFGDKLC 75
Db 40 EMLNLDLTIFFAQFVQATYKSVKMKDILTVEKSTGSEQGGCJENQJFAFLEIC 99

QY 76 TVATLRETYGEMADCCAKQEPERNECEFLQHKD-DNPMLPLRVRPEVDVMCTAFHDNEET 134
Db 100 HEKEISEKYG-LAOCSSQREERHNCFLAHKKAAPPSPFPQVAEPVTSCKAYEENEDMF 158

QY 135 LKKYLVEIARHPYPYAPELLFFAKRYKAAATECCQADKAACLLPKLDELREGKASSA 194
Db 159 MNRVIYEIARHPYAPYATILSLAHYKGIIPLCCKAENAVECPTQTSITLITKELRESSL 218

QY 195 XQRLKCSLQKGERAFKAMAVARLSORFPKAEVSKLVDLTLYKHTECHCDLLECA 254
Db 219 LNCHICAVMRNFGPTFAITVTKLQSFKSANTETIQKLVLDVAHIEBCCRNVLECL 278

QY 255 DDADLAKYICENQSISSKLECECKEPLEKSHCIAEVENDEMPADLPISLAADFVESKD 314
Db 279 QDGEKIMSVICSQQDILSSKIDCCCKLPILGQCIIHAENDGKPEGLSPNLNRFLEERD 338

QY 315 VCQNTAEAKDVLGMFLVEYARRHPDYSVLLLRKATYETITLXCCAAADPHCYAKVF 374
Db 339 FNQFSREKDLFMARFTVEYSRRHTKLAVPVVLVAKGYQELLEKCSQSENPCEQKGE 398

QY 375 DEFKPLVEEPQNLIKQCELEQGEYKFNALLVRVTKYPOVSTPTLVEVSENLCXVG 434
Db 399 BELEYIOESQALKRCGLFKQJGEYYLQNAFLVATYTKAAPQITPPPELMFTAFRMATAA 458
```

```
QY 435 SKCKCHPEAKEMPCAEDYLSVNLCLCVLHEKTPVSDRVTCKCTESLVNRR3PFCFSALEVC 494
Db 459 ATCCQLSDRQJACGEGAADLIIOQLCIRHEETPINPOVGQCCSSSYANRRPCFSSSLVVD 518
QY 495 ETVVPKEFNAETFTPHAD:CTLSEKERQIKKQTAALVELVGHKPKATKEOLKAVMDDFAA 554
Db 519 ETVYIPSPFSADKFIHKDLCQAQGVALQTMKQQLINLVKOKPO:TEBQLEAVIADFSL 578
QY 555 VEKCKKADDETCFAEBEGKLVAAASQALGL 585
Db 579 LEKCCQGEQEAEPFEEGPKLISKTRALGV 609

RESULT 7
Q8UMCS PRELIMINARY; PRT: 626 AA.
AC Q8UMCS;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)
DE Serum albumin precursor.
GN AaB.
OS Ambystoma maculatum (spotted salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OC NCBI_TaxID=43114;
RN [1]
RP SEQUENCE FROM N.A.
RA TissueLiver;
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217183; AAL56646.1; -
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN_3.
DR PROSITE; PS00212; ALBUMIN_3.
DR PROSITE; PS00212; ALBUMIN_3.
KW Signal.
FT SIGNAL.
FT CHAIN 1 24 POTENTIAL.
FT CHAIN 25 626 SERUM ALBUMIN.
SQ SEQUENCE 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;

Query Match 40.0%; Score 1242; DB 13; Length 626;
Best Local Similarity 40.1%; Pred. No. 2.8e-89;
Matches 237; Conservative 110; Mismatches 238; Indels 6; Gaps 3;

QY 1 DAHKSEVAHRFKDL---GEENFKALVLIAPAYLOCCPDEHVKLVNEVTEFAKTCVAD 56
Db 28 EGVNDVPHLIGDLIPMGVDNKGVLAAVYQMLPLCPYEEHLQVRVEDVMQIALCAKG 87

QY 57 ESABNCDKSLHTLFQDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDN-PNLRL 115
Db 88 ARHANCASPMTIILDELCKPENAEKYPFHQCECKEDPERHKGCFVEHKMANHEELTKY 147

QY 116 VRPEVDVMCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAAATECCQADKA 175
Db 148 VRPAPEIQCKQHAENRGPELLARYIIPMLAIGHPHMTIPAILGFAQRFDGIVSHCCKDVEIA 207

QY 176 A-CLLAPKLDELDEGKASSAKQRLKCSLQKGERAFKAMAVARLSORFPKAEVSKL 234
Db 208 GQCFNDKPEHKQEVYVYCALQKHCNYILQDFKERALTAYKAVQASOKFPLASFENVQII 267

QY 235 VTDLTAKVTECHCDLLECADRADLAKYICENQSISSKLECECKEPLEKSHCIAEVE 294
Db 268 VPDTVHLHTQCCGDMWACMLERMKLTAIKICEKDELATHLKECCDKPLERSACIIRLP 327

QY 295 NDEMPADLPISLAADFVESKDVCKNVAEKADVFLGMFLVEYARRHPDYSVLLLRKATY 354
Db 328 NDQKPADLSPKVPHYIDDPVECKLYTEGSDTDMGRFLYECARRHODYSPMLLRNGSGYE 387
```

QY	355	TTLEKCCAAADPHCEYAKVDFEFPKPLVEBPQNLIKONCELFQOLGEYKFKQNALVRYTKK	414
Db	388	EPLKCCAAEGHNECLAETESLKKEISSVTLKTCNGALDKLKSYLEFQNLIIKYYVAR	447
QY	415	VPOVSTPTLVEVRNLGKVGSKCKHPEAKRMPCADYLSVLNQLCVLHEKTPVSDRVT	474
Db	448	MPALSEQSLLRIITKETTIGKECHRPEDQOMTCSEGGIGVFGQICMKQKTPVNEKVA	507
QY	475	KCTESLVNRRPCFSALEVDVETVVPKFEFNAETFTPHADICTLSEKEROJKKOTALVELVK	534
Db	508	QCCSHLSSTQPCFSALPVDVETVVPPLSVASFNFNDELCTISEPQQSKKQVFJRLMK	567
QY	535	HKPEKATKEQLKAVMDFAAFVEKCKCKADOKETCFABEGKKLVAASQAALGL	595
Db	568	QYEHMTDEQLKTCVNVFVPMVDQCCCKADNHCEFALEGAKLIDACKAILAV	618
RESULT 8			
Q8MJ76			
ID	Q8MJ76	PRELIMINARY;	PRT; 610 AA.
AC	Q8MJ76;		
DT	01-OCT-2002 (T-EMBLrel. 22, Created)		
DT	01-OCT-2002 (T-EMBLrel. 22, Last sequence update)		
DT	01-MAR-2003 (T-EMBLrel. 23, Last annotation update)		
DE	Alpha-fetoprotein.		
OS	Sus scrofa (pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Kim J.G., Noneman D., Vallet J.L., Christenson R.K.;		
RT	"Mapping of the porcine alpha-fetoprotein (AFP) gene to SSC8.;"		
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF517770; AAM66710.1; -		
DR	InterPro; IPR000264; Serum_albumin.		
DR	Pfam; PF00273; transport_prot; 3.		
DR	PRINTS; PR00802; SERUMALBUMIN.		
DR	ProDom; PD002486; Serum_albumin; 1.		
DR	SMART; SM00103; ALBUMIN; 3.		
DR	PROSITE; PS00212; ALBUMIN; 2.		
SQ	SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;		
Query Match			
Best Local Similarity 39.3%; Score 1218.5; DB 6; Length 610;			
Matches 227; Conservative 117; Mismatches 225; Indels 3; Gaps 3;			
QY	16	BENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESANCKDSLHTLFGDKLC	75
Db	40	EMNLVDLTIFFPAQFVQEAITYKEVQMWKDVLTVEIKSTGSEQPACCLNQSVFLKEIC	99
QY	76	TVATLRETYGEMADCCAKQEPERNECFLOHKDNP-NLPRLVRPEVDVMTAFHNEETP	134
Db	100	HEEIEPEKYG-LSHCCSQSGEEHNCFLARKKAAPASIPFPQVPEVTSCKAYEENRELF	158
QY	135	LKKYLEIARRHPFYFAPELLFPFAKYKAAFTBCCQAADKAACLLKLDELDRDEGRASSA	194
Db	159	MRYIYIEIARRHPFLVAPILSLIAQYDKIIPCCKAENAVECFQTKAASITKEIRESSL	218
QY	195	KORLKCAASLOKGERAFKAWAVARLSQRFPKAFVSKLVTLDTKVHTECHGDLLECA	254
Db	219	LNQHMCTVMROFCARTFRAITVTKLSQKPPKANFTSIQKLV-LVAHIHEECCRGVLECL	278
QY	255	DDRADLAKYICENQDISSKLKECCBK-P-LLEKSHCIAEVENDENPADLPSLAAAFVSK	313
Db	279	QDAERVSVYVCSQDDTLSSKIAECCKLPTTLLELQGCIHAENDDKPEGLSPNLNRFLEGR	338
QY	314	DVCKYABAKDVLFGMLFLEYARRHPDYSVLLLRLLAKTYETTLKCCAAADPHCEYAKV	373
Db	339	DFNLSSREKDLSMARTYEYRRHPLAVPVLIRVAKGYQELLEKCSQSENPLECQDKG	398
QY	374	DFEFKPLVEBPQNLIKONCELFQOLGEYAKFQNALVRYTKKVPQVSTPTLVEVSRNLGKV	433

Db	399	EELEKYIOESQALAKRSCGLFQKGEVYLQNALVYTKKAPOLTPPELMALTKMMATT	458
QY	434	GSKCKHPEAKRMPCADYLSVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALV	493
Db	459	GAACCHJSEDQLACGGAADLIIGOLCIIRHEEPINPGVGQCCTSSYANRRPCFSSLLV	518
QY	494	DETVVPKFEFNAETFTPHADICTLSEKEROIKKOTALVELVKHKPKATKEQLKAVMDFAA	553
Db	519	DETVVPKFEFNAETFTPHADICTLSEKEROIKKOTALVELVKHKPKATKEQLKAVMDFAA	578
QY	554	FVEKCKKADDKETCFABEGKKLVAASQAALGL	585
Db	579	LLEKCCQCEQEVCFABEGPALISKTRASLGV	610
RESULT 9			
Q8JWC6			
ID	Q8JWC6	PRELIMINARY;	PRT; 624 AA.
AC	Q8JWC6;		
DT	01-MAR-2002 (T-EMBLrel. 20, Created)		
DT	01-MAR-2002 (T-EMBLrel. 20, Last sequence update)		
DT	01-MAR-2003 (T-EMBLrel. 23, Last annotation update)		
DE	Serum albumin precursor.		
GN	ALB.		
OS	Ambystoma texanum (Smallmouth salamander).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomidae;		
OC	Ambystoma.		
OX	NCBI_TaxID=8104;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;		
RT	"Serum albumin of the mole salamanders Ambystoma maculatum and		
RL	Ambystoma texanum.;"		
DR	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF217182; AAL56645.1; -		
DR	InterPro; IPR000264; Serum_albumin.		
DR	Pfam; PF00273; transport_prot; 3.		
DR	PRINTS; PR00802; SERUMALBUMIN.		
DR	ProDom; PD002486; Serum_albumin; 1.		
DR	SMART; SM00103; ALBUMIN; 3.		
DR	PROSITE; PS00212; ALBUMIN; 1.		
KW	Signal.		
FT	SIGNAL	1 24	POTENTIAL.
FT	CHAIN	25 624	SERUM ALBUMIN.
SQ	SEQUENCE 624 AA; 70321 MW; DE08533BF4953EF7 CRC64;		
Query Match			
Best Local Similarity 35.0%; Score 1087; DB 13; Length 624;			
Matches 214; Conservative 102; Mismatches 240; Indels 12; Gaps 5;			
QY	14	LGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESANCKDSLHTLFGDK	73
Db	45	IGVEHAKALAMALFSQMSLCKPHEHQVQVRVNVMDIADLCRGAHGDCGKSVMTIINE	104
QY	74	LCTVATLRETYGEMADCCAKQEPERNECFLOHKDNP-NLPRLVRPEVDVMTAFHNEE	132
Db	105	ICKTPEPEKYPFHGGCKCKEDPERHKCF-EHKSTDPKERTYEVYKPSFQICKSHAENRD	164
QY	133	TFLKKYLEIARRHPFYFAPELLFPFAKYKAAFTBCCQAADKAACLLKLDELDRDEGKA	191
Db	165	EFLGHYIHKVASSHTTMYPPALLSFTLHFDGIVSHCKCKDEATVGCCLSEKMPAKHEVEH	224
QY	192	SSAQORLKCAASLOKGERAFKAWAVARLSQRFPKAFVSKLVTLDTKVHTECHGDL	251
Db	225	VCAVQKNCYILQNFERNALRASKAHACSKFPHASFENVQRLTDGIVHLHOTCGGDM	284
QY	252	ECADDRADLAKYICENQDISSKLKECCBK-P-LLEKSHCIAEVENDENPADLPSLAAAFV	311
Db	285	ACMAERMKLTJTQCEK-----KCCQKPVLSERSEIVLPLNDEKPADLSPEVRYFD	336
QY	312	SKVCKKNYAEKDVFLGMFLFLEYARRHPDYSVLLLRLLAKTYETTLKCCAAADPHCEYA	371

```

Db      337 DPEVCKRKEGDAPYGRFLCYAKIPEHSAENLRIASGLEKAYKTCGAGAHNECIA 336
QY      372 KVDFEFKPLVEEPQNLIKONCELFQGLGEYKFQNALLVRYTKKVPQVSTPTLVEVSNL 431
Db      397 KEEETLRHEIEASKTKLTTCGALKGLPYHFQINIMIVRYTGILPRSSDAFLVYTKLT 456
QY      432 KUGSCCKHPKAMPKPCADYLSVVLNQLCVLHEKTPV-SDRVTKCTESLVNRRPCFSA 490
Db      457 NIGQCKCKLPDQOMPCSEGGLGVFAOIC-ONQKTPFENEKJAHCKDLSLFT-PCFAA 515
QY      491 LEVDYVYVPEFNAETFTFHADICTLSEKERQIKKOTALVELVHKPKATKEQLKAYMOD 550
Db      516 LTJVDYVYPAVTAESFNFNEDEFTPEADLQAKKQFLMLHVRTHPKITDEQVKTSX 575
QY      551 PAAVFEKCKKADDETCFASEGKLVAA 578
Db      576 FLAMGQCKKADQORNECPATEGAKLVEA 603

RESULT 10
Q8BK65 PRELIMINARY; PRT; 605 AA.
AC Q8BK65;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK076053; BAC36150.1; -.
SQ SEQUENCE 605 AA; 67322 MW; 049B7A4A8B01E4B CRC64;

Query Match 34.9%; Score 1082; DB 11; Length 605;
Best Local Similarity 35.7%; Pred. No. 1e-76;
Matches 204; Conservative 120; Mismatches 241; Indels 6; Gaps 3;

QY 16 BENFKALVLIAPAQYLOQCPFEDHVKLNEVTEFAKTCVADESACNCKSLHTLFGDKLC 75
Db 40 EKNVLSIATITFTQVPEATEEE---VNKMTSDVLAAMKXNSGGCLESQSVFLDEIC 95
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDDNP-NLPRLVRPEVYVWCTAFHNEET 134
Db 96 HETELSNKYG-LSGCCSGSVERHOCLLARKKTAPASVPPFPPEPASCKAHEENRAVF 154
QY 135 LKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLPKLDELRDEGKASSA 194
Db 96 HETELSNKYG-LSGCCSGSVERHOCLLARKKTAPASVPPFPPEPASCKAHEENRAVF 154
QY 135 LKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLPKLDELRDEGKASSA 194
Db 155 MNRFIYVSRNPFMYAPAILSLAAQYDKVVLACCADNKECFQTKRASAKELREGSM 214
QY 195 KORLKCASLOKGEGERAFKAWAVARLSQRPKAPAEVSKLVTDLTKVHTECHGDLLSCA 254
Db 215 LNEHVCVIRKFGSRNLQATTIIKLSQKLTEANFTEIQKALDVAIHIEECQCGNSLECL 274
QY 255 DORADLAKYICENQDISSKLEKCEKPLEKSHCIAEVENDEMPADLPSLAADFVESKD 314
Db 275 QDGEKVMYICSQQNLSSKIAECCKLPMIQLGFCIIHAENGVRPEGLSNPSQFLGDRN 334
QY 315 VCNVYAEAKDVFGLMFLEYARRHPDYSVLLLRLLAKTYETTLKCCAAADPHECYAKVF 374
Db 275 QDGEKVMYICSQQNLSSKIAECCKLPMIQLGFCIIHAENGVRPEGLSNPSQFLGDRN 334
QY 315 VCNVYAEAKDVFGLMFLEYARRHPDYSVLLLRLLAKTYETTLKCCAAADPHECYAKVF 374
Db 335 FAQFSSEKIMFMAFLHEYSRTHPNLPVSVILRIAKTYQEIILEKCSQSGNLPGCCDNLE 394
QY 375 DEFKPLVEEPQNLIKONCELFQGLGEYKFQNALLVRYTKKVPQVSTPTLVEVSNLKG 434
Db 335 FAQFSSEKIMFMAFLHEYSRTHPNLPVSVILRIAKTYQEIILEKCSQSGNLPGCCDNLE 394
QY 375 DEFKPLVEEPQNLIKONCELFQGLGEYKFQNALLVRYTKKVPQVSTPTLVEVSNLKG 434

```

```

Db      395 EELQKHIEESQALSQSCALYQTLGDYKQLNQLFLGYTRKAPOLTSBELIDLTGMWSIA 454
QY      435 SKCKHPKAMPKPCADYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVC 494
Db      455 STCCQLSEBKSGCGEGNADIFIGHLCIRNEASPVNSGISHCCNSSYSNRRLCTTSFLRD 514
QY      495 ERYVPEKPENAETFTFHADICTLSEKERQIKKOTALVELVHKPKATKEQLKAYMCDFAAF 554
Db      515 EYIAPPPFSEDKPIPHKOLCQAHKALQTMKQELINLVKOKPELTBEQJAAVTAADFSG 574
QY      555 VEKCKKADDETCFASEGKLVAAASQAALGL 585
Db      575 LEKCKCAQOQEVCFTEEGPKLISKTRDALGV 605

RESULT 11
Q8BK56 PRELIMINARY; PRT; 605 AA.
AC Q8BK56;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK076197; BAC36249.1; -.
SQ SEQUENCE 605 AA; 67409 MW; CF3509A9EC146110 CRC64;

Query Match 34.7%; Score 1076; DB 11; Length 605;
Best Local Similarity 35.6%; Pred. No. 3.1e-76;
Matches 203; Conservative 120; Mismatches 242; Indels 6; Gaps 3;

QY 16 BENFKALVLIAPAQYLOQCPFEDHVKLNEVTEFAKTCVADESACNCKSLHTLFGDKLC 75
Db 40 EKNVLSIATITFTQVPEATEEE---VNKMTSDVLAAMKXNSGGCLESQSVFLDEIC 95
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDDNP-NLPRLVRPEVYVWCTAFHNEET 134
Db 96 HETELSNKYG-LSGCCSGSVERHOCLLARKKTAPASVPPFPPEPASCKAHEENRAVF 154
QY 135 LKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLPKLDELRDEGKASSA 194
Db 155 MNRFIYVSRNPFMYAPAILSLAAQYDKVVLACCADNKECFQTKRASAKELREGSM 214
QY 195 KORLKCASLOKGEGERAFKAWAVARLSQRPKAPAEVSKLVTDLTKVHTECHGDLLSCA 254
Db 215 LNEHVCVIRKFGSRNLQATTIIKLSQKLTEANFTEIQKALDVAIHIEECQCGNSLECL 274
QY 255 DORADLAKYICENQDISSKLEKCEKPLEKSHCIAEVENDEMPADLPSLAADFVESKD 314
Db 275 QDGEKVMYICSQQNLSSKIAECCKLPMIQLGFCIIHAENGVRPEGLSNPSQFLGDRN 334
QY 315 VCNVYAEAKDVFGLMFLEYARRHPDYSVLLLRLLAKTYETTLKCCAAADPHECYAKVF 374
Db 335 FAQFSSEKIMFMAFLHEYSRTHPNLPVSVILRIAKTYQEIILEKCSQSGNLPGCCDNLE 394
QY 375 DEFKPLVEEPQNLIKONCELFQGLGEYKFQNALLVRYTKKVPQVSTPTLVEVSNLKG 434
Db 395 EELQKHIEESQALSQSCALYQTLGDYKQLNQLFLGYTRKAPOLTSBELIDLTGMWSIA 454
QY      435 SKCKHPKAMPKPCADYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVD 494

```

Db	455	STCCOLSEKWSGCCGGMADIFIGHLCIRNEASPVNSG:SHCCNSYSNRRLCIT-SFLRD	514
QY	495	ETYPKEFNAETTFHADICTLSEKERO:KKOTALVELNKHGPKATKEQLKAVMDQFAAF	554
Db	515	ETIAPPPSDFKFIHKDLQACGKALQTMKQELLINLVKQKFLTEBEOAAVTAADFSG	574
QY	555	VEKCKADDKETCFABEGKLVAAQAALGL	585
Db	575	LEKCKCAQDOEVCFTTEEGPKL:SKTRDALGV	605
RESULT 12			
Q8JIA7	PRELIMINARY; PRT; 400 AA.		
AC	Q8JIA7;		
DT	01-OCT-2002 (TREMBlrel. 22, Created)		
DT	01-OCT-2002 (TREMBlrel. 22, Last sequence update)		
DE	01-MAR-2003 (TREMBlrel. 23, Last annotation update)		
DE	A/B over-sized serum albumin (Fragment).		
OS	Sphenodon punctatus (Hatteria) (Tuatara).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.		
OX	NCBI_TaxID=8508;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;		
RT	"Partial mRNA sequence for tuatara A/B serum albumin."		
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF375973; AAM461C6.1; --		
DR	InterPro; IPR000264; Serum_albumin.		
DR	Pfam; PF00273; transport_prot; 2.		
DR	PRINTS; PR00802; SERUMALBUMIN.		
DR	ProDom; PD002486; Serum_albumin; 1.		
DR	SMART; SM00103; ALBUMIN; 2.		
DR	PROSITE; PS00212; ALBUMIN; 2.		
FT	NON_TER		
FT	1		
SQ	SEQUENCE 400 AA; 45715 MW; 8DE20609657CF753 CRC64;		
Query Match 33.7%; Score 1045; DB 13; Length 400;			
Best Local Similarity 47.7%; Pred. No. 5.1e-74;			
Matches 186; Conservative 80; Mismatches 124; Indels 0; Gaps 0;			
QY	196	QRLKASLCKFGGERAFKAWARLSORFPKAEVSKLVTLTKVHTCCGDLLECAD	255
Db	3	EKHSCGFLKSGGERAFQANKLVLSQKFPKPFEEHKVLTATLKQECCHGQIECJD	62
QY	256	DRADLAKYTCENQDSISSKLKSCCEKPLEKSHCIAEVENDEMPADLPISLAADFVESKDV	315
Db	63	DRVEVMAYICSKQAVFSSIKDCCEKPIVDREVCIQAOLDEKPADLPISAGYIESTEV	122
QY	316	CKNYAEAKDVLGMFLVEYARHPDYSVLLRLAKTYETTLKCCAAADPHECYAKVD	375
Db	123	KHYBEGKOVFLAHFVYVSRHRHPSSQMLRTGQYQDTLDKCKTENPPECYKAGE	192
QY	376	BFKPLVEEPQNLIKONCELFQEGYKFNALLVRYTKKVPQVSTPTLVEVSRNLGKVS	435
Db	183	ELARHIQESQELKTHCSFYTSQKDPFKQVLYRYTKOMPQLPAELLIISKKLTGVGV	242
QY	436	KCKHIPEAKRMPCAEDYLSVNLQLCVJHEKTPVSDRVTKCTTESLVNRRPFCFSALEVD	495
Db	243	KCCPLSEDKRLSCSEKLSMWLFECIROHEASPVNNHVTHCCTDSYSEMRPCTKLGVD	302
QY	496	TYVPKEFNAETTFHADICTLSEKROIKKOTALVELVKHPKATKEQLKAVMDQFAAF	555
Db	303	SYVPPEFCFSTFLFDQLCTAPEEARLKKQLTFLVKLIQLKPFQIEDEQLKLVTDYHAME	362
QY	556	EKCKADDKETCFABEGKLVAAQAALGL	585
Db	363	EKCCQAEKQECFSTEGEKLTOEGKALLGV	392

RESULT 13

RESULT 13

Q9YGH6	PRELIMINARY; PRT; 603 AA.	
ID	Q9YGH6;	
AC	Q9YGH6;	
DT	01-MAY-1999 (TREMBlrel. 10, Created)	
DT	01-MAY-1999 (TREMBlrel. 10, Last sequence update)	
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)	
DE	Serum albumin precursor (Fragment).	
GN	ALB.	
OS	Rana shqiperica.	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.	
OX	NCBI_TaxID=44326;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	STRAIN=Bushat; TISSUE=Liver;	
RA	Uzzell T., Hotz H.;	
RT	"Albumin cDNA sequence of Rana shqiperica: evolutionary changes in frog albumins."	
RL	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; U40452; AAD09358.1; --	
DR	HSSP; P02768; 1E7B.	
DR	InterPro; IPR000264; Serum_albumin.	
DR	Pfam; PF00273; transport_prot; 3.	
DR	PRINTS; PR00802; SERUMALBUMIN	
DR	ProDom; PD002486; Serum_albumin; 1.	
DR	SMART; SM00103; ALBUMIN; 3.	
DR	PROSITE; PS00212; ALBUMIN; 3.	
KW	Signal.	
FT	NON_TER	1
FT	SIGNAL	<1 23
FT	CHAIN	24 603
SQ	SEQUENCE 603 AA; 69293 MW; 340D3723FA010C99 CRC64;	
Query Match 30.8%; Score 955; DB 13; Length 603;		
Best Local Similarity 33.6%; Pred. No. 9.9e-67;		
Matches 190; Conservative 119; Mismatches 241; Indels 16; Gaps 3;		
QY	12	KDLGEENFKALVLIAPQYILQCCFFEDHVKLVNVEVTEFAKTCVADESAENCDKSLHTLFG 71
DB	37	KAVGKPAVEKLVVMVAQDFEKCSDHEHLKVQAKIIKAVNCEKHPKEAECKKPAIELYH 96
QY	72	DKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPLVR-----PEVDVMT 125
DB	97	DIVCKEEDIDQLYPWTTECCGKABERTKCFYEHRE-----VRVEYKIPNIESCK 148
QY	126	AFHNEETFLKYLIEIARHPYFAPELLFFAKRYKAAATECCQAAADKAACLPKLDEL 185
DB	149	EKHEHPORAFSYLLSNIAKRSKLYPPAVLGFATQYNEITTECCAAEDKAKCFGERMPQV 208
QY	186	RDEGKASSAKORLKACSLQKFGBERAFKAWARLSORFPKAEVSKLVTLTKVHTEC 245
DB	209	KKLTNYLEDKHKQKRVLKBPFRVSQALTLVQVSQFQGNKYDDVEKVTIEIAHLNEDC 268
QY	246	CHGDLLECAADRADLAKYTCENQDSISSKLKSCCEKPLEKSHCIAEVENDEMPADLP 305
DB	269	CKGDAVECMIERMEATEHICLAEKLSKLSDCCKAGVLETPCILALPNEE--PDLPIE 326
QY	306	AADFVESKDVCKNYAEAKDVLGMFLVEYARHPDYSVLLRLAKTYETTLKCCAAAD 365
DB	327	LKEYEDEVHCENYQKQKRYLAHFTHDYSRSHOESSPOSCLRVSRGFEMLEKCCASAN 386
QY	366	PHCYAKVDFBPKPLVEEPQNLIKONCELFQEGYKFNALLVRYTKKVPQVSTPTLVE 425
DB	387	SAECLDKAPKLEAALKENEISIKQNGCALEKLGFNDFYQLLVRYFGKMPQVTAQLVE 446
QY	426	VSRNLGKVGSKCKHPKAEKRMPCAEYLSVNLQLCVJHEKTPVSDRVTKCTTESLVNRR 485
DB	447	LTGRMAKIGVYCCGLPDNKKQPCAEKLDILLGEMCEREKKTFINDNVHCCVDVYANRR 506
QY	486	PCFSALEVDETYVPKEFNAETTFTHADICTLSEKROIKKQATLVELVKHPKATKEQLK 545
DB	507	PCFTKLGFPYANYEAPVWVDESKLHFTADMCKGSADDGLKTLVLLVFEFLKMKPTCGEKL 566

Qy 532 LVKHKPKATKEQLKAVMDDFAAFVEKCCKADDDKETCFPAEKGKXLVAASQAAL 583
Db 354 VVKCKPAITHEQLKAVITDFYGVVEKCCGCHGENHEACFLAEGPOLVORTQAAL 405

Search completed: October 27, 2003, 15:07:11
Job time : 110 secs